Imperial College London

Department of Electrical and Electronic Engineering

Modelling Brexit

Supervisor:
Dr Giordano Scarciotti

Author:
Victor Sanchez

Second marker:
Dr David Angeli

A report submitted for the degree of

BEng Electronics and Information Engineering

June 20, 2019

Abstract

This project aims at designing a comprehensive and realistic opinion dynamics model to study the evolution of opinion in a population in the event of a two-choices referendum. The model was implemented following the case-study of the Brexit referendum which allowed for parameter tuning and performance evaluation. The main aspect of the project consists in implementing a basic prototype inspired by state of the art models and improve it by adding on features to increase complexity and include components which typically influence an individual's decision making process. Using MATLAB, a population of agents representing the United Kingdom population was simulated, each agent characterized by age, employment, salary, education and stubbornness which influence initial opinion and its shaping when interacting with others. How certain events impact individuals and population's opinion was also studied as a meaningful component of opinion shaping in real-life scenario. The evaluation procedure consisted in evaluating how well the model can reproduce the EU Referendum at initialization, before assessing how close the opinion evolution produced by the model fits the polls conducted since the referendum took place. Future improvements to this project can consist in refining parameters and taking into account more aspects of public opinion shaping such as media (internet, newspaper, television, social network), employed agents (used by political parties to influence on population) and ethnic origins.

Acknowledgements

I would like to take the opportunity to thank the following people who have contributed directly or indirectly to the production of this project and report:

• **Dr Giordano Scarciotti** for his availability, support and advice thoughout this term.

• **Stevan Vukmirovic** for proof-reading this report and providing a harsh but more than appreciated critical view of my approach to this project.

• My girlfriend **Amélie** for her continuous support throughout this third year.

• All my friends and the people who have made these three years at Imperial College the most memorable of my life.

Contents

1 Introduction **5**

1.1 Motivation . . . 5

1.1.1 Brexit referendum . . . 5

1.1.2 Opinion dynamics modelling . . . 6

1.2 Objectives . . . 6

1.2.1 Features . . . 6

1.2.2 Evaluation . . . 7

1.3 Structure of the report . . . 7

2 Literature Review **8**

2.1 Cognitive biases . . . 8

2.1.1 Confirmation bias . . . 8

2.1.2 Anchoring bias . . . 8

2.1.3 Continued Influence Effect . . . 9

2.1.4 Groupthink . . . 9

2.2 Discrete opinion modelling . . . 9

2.2.1 Ising Model and Strikes . . . 9

2.2.2 Voter model: invasive approach . . . 10

2.2.3 Group majority . . . 10

2.2.4 Hierarchy majority . . . 11

2.2.5 United we Stand Divided we Fall (Sznajd) . . . 11

2.2.6 Comments . . . 12

2.3 Continuous opinion modelling . . . 13

2.3.1 Deffuant model . . . 13

2.3.2 DeGroot: individual influence in a group . . . 14

2.3.3 Friedkin-Johnsen Model . . . 15

2.3.4 Comments . . . 15

2.4 Other literature of interest . . . 16

2.4.1 Deadlock and consensus reaching: the case of bees . . . 16

2.5 Further literature (for the interested reader) . . . 17

2.5.1 Further reading on discrete opinion modelling . . . 17

2.5.2 Further reading on discrete continuous opinion modelling . . . 17

2.5.3 Further reading on other topics . . . 18

3 Analysis and Design **19**

3.1 High level view . . . 19

3.2 Implementation . . . 21

3.2.1 Simple models - 3-dimensional array implementations . . . 21
3.2.2 Later models - object-oriented approach . . . 21
3.3 Development protocol - incremental and iterative approach . . . 21
3.4 Experimenting Protocol . . . 22
3.5 Metrics . . . 22
3.5.1 Statistical estimators . . . 22
3.5.2 Referendums . . . 23

4 Analysis and design **24**
4.1 Group mean opinion . . . 24
4.2 Group Deffuant interactions . . . 25
4.3 Introducing stubbornness . . . 29
4.4 Introducing age and initial distribution bias . . . 31
4.5 Normal biased initial opinion distribution . . . 33
4.6 Re-calibrating parameters to obtain realistic results under normal distribution . . . 35
4.6.1 Changing group size: results . . . 35
4.6.2 Candidate stubbornness functions: results . . . 36
4.7 Convergence rate - decrease over time . . . 37
4.8 Convergence rate - agent assigned . . . 39
4.9 Adding to initial opinion distribution: gender, household income, education and political attention . . . 40
4.9.1 Discussion . . . 41
4.10 Indecisiveness and turnout . . . 41
4.11 Events . . . 43
4.11.1 Simulation 1: common shift . . . 44
4.11.2 Simulation 2: agent-specific shift . . . 45
4.11.3 Simulation 3: multiple random events . . . 46

5 Testing **48**
5.1 Evaluation . . . 48
5.1.1 Modelling Great Britain . . . 48
5.1.2 Modelling West Midlands . . . 51
5.1.3 Modelling Scotland: model limits . . . 52
5.1.4 Modelling Greater London . . . 54

6 Conclusion and future work **56**
6.1 Discussion and self-critique . . . 56
6.1.1 Discussing the model . . . 56
6.1.2 Methodology . . . 56
6.1.3 Assumptions and improvements . . . 56
6.2 Future work: other interesting ideas . . . 57
6.2.1 Socio-demographics . . . 57
6.2.2 Media . . . 57
6.2.3 Employed agents . . . 58
6.2.4 Discrete opinion model . . . 58

Bibliography **66**

A Preliminary experiment - Deffuant model **67**
A.1 Purpose 67
A.2 Running documented experiments 67
A.2.1 Experimenting 67

Chapter 1

Introduction

1.1 Motivation

1.1.1 Brexit referendum

Context

On the 20^{th} of January 2016, United Kingdom Prime Minister David Cameron announced that a national referendum to answer the question "Should the United Kingdom remain a member of the European Union or leave the European Union?" would be held on the 23^{rd} of June 2016 [1]. The two possible answers were as follows:

1. "Remain a member of the European Union"
2. "Leave the European Union"

which will be referred to as remain and leave respectively. The official results were announced on the 24^{th} of June at 7:20 BST:

	Votes	%
Remain	16,141,241	48.11%
Leave	17,410,742	51.89%
Void	25,359	N/A
Turnout	46,500,001	72.21%

Table 1.1: Official results to the EU Referendum held on 23^{rd} of June 2016 [2]

A majority of the population opted for leaving the European Union.

Significance

The outcome of the Referendum has a strong cultural and political significance as it is the first time in the history of the European Union that article 50, stating "Any Member State may decide to withdraw from the Union in accordance with its own constitutional requirements" was summoned. Speculations by experts on the consequences for the Union and United Kingdom range from major recessions to mere flick to the economy, and negotiations regarding which deal will be adopted have been lasting for the past 3 years. While political and economical implications are still vague and hard to predict, it is not the topic of this project, which will instead focus on the opinion dynamics behind the referendum which led to this very decision.

1.1.2 Opinion dynamics modelling

Opinion modelling is a field of mathematical modelling which aims at achieving a description of how a group of individuals form their respective and collective opinion on a particular topic. Such models are then used to explain certain situations, predict how a certain opinion will evolve over time in a given population or simply understand why a certain population reached a decision. Comprehensive description and discussion of state-of-the-art models designed in the past 70 years is available in Section 2. The main challenges posed to opinion models are to clearly understand the forces that drive opinion evolution in a given population and take into account all the parameters involved. As in all mathematical modelling domains, opinion modelling faces the trade-off between having a simple model allowing for interpretation and rigorous analysis at the detriment of realism, while more comprehensive models are harder to analyze but can encompass more factors encountered in real life scenarios.

Opinion dynamics and the Brexit referendum

The EU Referendum, between its announcement and conduct, provides us with a short time-frame in which many polls have allowed to follow the population's opinion dynamics. Very accurate data on opinion distribution was collected on the days following the referendum, providing with substantial insight into the population's opinion at a given time, followed by regular polls which are still conducted today, providing us with great details on the population's opinion dynamics.
This project aims at designing a comprehensive and realistic opinion dynamics model to study the evolution of opinion in a population in the event of a two-choices referendum.The EU Referendum will act as a real-life case study of opinion dynamics, which we will use to shape our model's initial opinion distribution, tune opinion evolution parameters and assess the quality of this very model.

1.2 Objectives

1.2.1 Features

The presented model will attempt to favour comprehensiveness over simplicity, comprising the following features:

Initial opinion distribution and geography

Opinion distribution regarding the topic of Brexit has been shown to be closely correlated with socio-demographic factors such as age, education and employment. Individual's opinion within the population will be dependent upon these factors, which will be distributed depending on which region is modelled following Government's official statistics.

Population fragmentation and social mobility

The modelled populations are not a single mass of individuals, but rather a collection of sub-groups which represent social-circles to which individuals belong and within which debates and interactions will occur, driving dynamically opinion evolution. To account for the fact that agents tend to change their social environment from time to time for various reasons (changing jobs, moving to a new city or neighbourhood, picking up a new activity), social mobility will be accounted for by having agents randomly change social groups on a regular basis.

Cognitive biases shaping interactions

Individuals do not all interact similarly, primarily because of cognitive bias which vary from an agent to another. Interactions are therefore shaped according to parameters which will either precipitate or prevent opinion updates within social-groups, specifically to each agent.

Events

Unusual, exceptional or shocking event have been shown to have an impact on individual's opinion on related topics. We therefore incorporated into the model such events occurring, allowing for the study of their impact on opinion shaping in a population.

More details on motivations and implementation of these features can be found in Section 4.

1.2.2 Evaluation

The model, developed on a general frame with parameters representing the global UK population will be evaluated on this same population and two specific sub-regions. The model will be tuned to model opinion distribution and population characteristics of these three regions on the days following the referendum. Assessment will be done at two stages: initial opinion distribution should correspond to the referendum results for the modelled region, and the resulting opinion evolution should match the recent polls conducted in the region. The model will finally be ran in an attempt to forecast the opinion changes of a fourth region.

1.3 Structure of the report

Chapter 2 introduces relevant notions on cognitive biases necessary for understanding their relevance in the model, as well as state-of-the-art models in opinion dynamics, providing the reader with context and information on relevant works which have acted as stepping-stones for designing the final model presented in Chapter 3. Chapter 3 will also introduce the experimental protocol used to build the model. Gradual improvements and the adopted detailed approach are presented in Chapter 4. Assessment of the final model is described in Chapter 5 which comprises the evaluation method, results and discussion. Final conclusion and sources of improvements are discussed in Chapter 6.

Chapter 2

Literature Review

2.1 Cognitive biases

We here present cognitive biases: the tendency of our brain to introduce bias when processing information - also called heuristics - or when making a decision. These cognitive biases are of importance when trying to model opinion dynamics as they are a core reason for why human agents do not necessarily behave in a rational way and tend to take decisions and forge their opinion in a unexpected way. Most models (as described in Section 2.2 and 2.3) try to incorporate these either by accounting for them as noise, or by introducing a specific parameter to model their influence.

2.1.1 Confirmation bias

According to Nickerson [3], confirmation bias "connotes the seeking or interpreting of evidence in ways that are partial to existing belief", that is agents will tend to seek information going in the sense of their opinion or interpret available information as such. In the context of decision making, we can assume agents will tend to adopt the views or be influenced by neighbours who share their beliefs and point of view rather than those with diverging opinion [4].

2.1.2 Anchoring bias

The anchoring bias was first introduced by Tversky and Kahneman [5] in order to show how economic agents could behave in an irrational manner when assessing the probability of a particular event. Agents subject to this bias tend to establish their estimates with a strong bias towards an initially presented value. A good illustration of the principle is the experiment used by the two psychologists to demonstrate their point. Two groups of subjects were asked to estimate under very limited time (5 seconds) a numerical expression, given to each group in different order. Group 1 had to estimate $8\times7\times6\times5\times4\times3\times2\times1$. Group 2 was given the expression $1\times1\times3\times4\times5\times6\times7\times8\times$. The paper states the following result: group 1 (descending sequence) gave a median answer of $2,250$ while group 2 (ascending sequence) gave a median result of 512. Researches were expecting an underestimate for both group (the correct answer is $40,320$) and concluded that given the very limited time frame, most subjects only had time to compute the very first products of the expression before giving their estimate, suggesting they had to rely on a higher result for group 1 while group 2 would rely on a lower result, thus explaining the difference in estimates between both groups. A comprehensive literature review of the topic by Furnham and Boo [6] shows how further studies have suggested that the anchoring bias could also influence individuals in shaping their opinion.

In a political situation, agents will therefore tend to judge candidates and programs based on their initial opinion (what party they already follow) which acts as a reference point.

2.1.3 Continued Influence Effect

Anchoring bias is furthermore reinforced by what is called the Continued Influence Effect, that is the tendency observed in certain individuals to keep on believing previously learned information even it after it has been shown to be erroneous and corrected. That is, the anchoring effect can persist and influence decisions and opinion formation even after learning that the original opinion is false [7]

2.1.4 Groupthink

Groupthink, as defined by Janis [8], is a term inspired from Orwell's "doublethink" which describes the dynamics within a group which causes members to conform to the majority opinion and even believe it. The effect was well demonstrated in 1951 by Solomon Asch in what is known as the Asch experiment [9]: a group of individuals was placed in an elevator, standing facing a specific direction (facing the door or turning their back to it). A subject was then introduced who would get on the elevator once every individuals was already in place. The results show that in a vast majority of cases, subjects adopt the same standing position as all the individuals in the group, facing the same direction. This result has been shown again in various experiments. In the case of opinion modelling, some individuals will have a strong tendency to have their opinion biases if they happen to be in a group where a different opinion is in majority, in order to conform (whether consciously or not).

The three following sections aim at presenting, reviewing and commenting on some important models of opinion dynamics introduced in the past 60 years in the light of this project. They will be considered distinctly as follows: discrete opinion modelling, continuous opinion modelling and miscellaneous. A last section will briefly provide further readings which will not contribute to the project but are still relevant for any reader curious to learn more about the topic.

2.2 Discrete opinion modelling

Discrete opinion dynamics models aim at studying a population of agents whose opinion can be represented with the help of a discrete variable. Each agent therefore holds an integer valued opinion which, depending on the situation, can represent their answer to a yes-no question, which candidate they plan on voting for at the next election... This section presents several discrete opinion models, most of which are based on a statistical physics framework.

2.2.1 Ising Model and Strikes

Ising model

Ernst Ising [10] established in 1925 a statistical physics model to describe an ensemble of atomic spins aligned in a lattice from which emerges a macroscopic magnetic moment (also called ferromagnetism). Spins can be in either of two states: $\{UP, DOWN\}$ or $\{+1, -1\}$. The total energy of a given spin configuration of n spins $\sigma \in \{-1, +1\}^n$ is:

$$H(\sigma) = -J_{ij} \sum_{<i,j>} \sigma_i \sigma_j$$

The sum is computed over all pairs of atomic spins. J corresponds to the interaction strength.

Application to strikes

In 1985, Serge Galam [11] had the idea of using the Ising model to describe the collective state of a cohort of agents which could be in either of two states: on strike or working. Each agent i is therefore assigned a spin $\tilde{\mu}_i$, $i = 1, ..., N$ with N total count of agents. An agent with $\tilde{\mu}_i = 0$ is on strike while an agent with $\tilde{\mu}_i = 1$ is working. The individual's state is then normalized as $\mu_i = 2(\tilde{\mu}_i - \frac{1}{2})$ which gives an Ising variable for agent's state:

$$\mu_i = \begin{cases} 1, \text{ if agent is working} \\ -1, \text{ if agent is on strike} \end{cases}$$

Interaction strength between two different agents i and j, $i \neq j$ is taken into account via the coupling variable $J_{ij} > 0$ which allows to account for the fact that individuals tend to reproduce their neighbour's opinion. Galam et al. then add an "external field" $H = W - E$ where W is the wage (common to all workers) and E the agent's expectations in term of wage. Linear coupling $H\mu_i$ allows to model (dis)satisfaction as $W > E$ is an agent satisfied with his salary while $W < E$ is an agent expecting a raise. Summing agent interaction and satisfaction we obtain:

$$\mathcal{H} = -\sum_{(i,j)} J_{ij} \mu_i \mu_j - H \sum_{i=1}^{N} \mu_i$$

2.2.2 Voter model: invasive approach

Initially named "Model for Spatial Conflict", the model was introduced by Clifford and Sudbury [12] in 1973 as a way to model interaction between two species competing over some territory. It was later renamed the voter model for its adequacy to model a population of voters holding an opinion evolving according to their neighbour's voting intention. Each individual holds an opinion quantified by a binary variable : $\{-1, +1\}$. At every time step, a random individual is selected and will grant its opinion to one of his neighbouring agents (picked randomly as well).

2.2.3 Group majority

This model was introduced by Serge Galam [13] to study opinion dynamics in a population in the context of everyday social interactions (dining in a restaurant, drinking with friends in a pub, lunch with colleagues). A finite population of N individuals is considered, which have to decide whether or not to accept a decision. At initial time t_0, $N_+(t_0)$ individuals support the proposition while $N_-(t_0) = N - N_+(t_0)$ are against it. At each time iteration $(t + 1)$ individuals randomly meet in small groups where each involved individuals will adopt the group majority opinion.

2.2.4 Hierarchy majority

Also introduced by Serge Galam [14], this model does not study opinion dynamics in the context of social interactions but rather in the case of subgroups of individuals voting to elect a representative. A finite population of N individuals is considered where two political opinions A and B are distributed with respective probability p_0 and $q_0 = 1 - p_0$. We form $\frac{N}{r}$ cells of r individuals which constitute the level 0 of the hierarchy. Each cell elects a representative following the voting rule $R_r(p_0)$ for opinion A or B which depends on the initial composition of the group. The elected representative of a group at the first level is therefore either of opinion A with probability $p_1 = R_r(p_0)$ or of opinion B with probability $q_1 = 1 - p_1$. Electing representatives for each group of level 0 constitutes a new group of elected individuals which are the first level of the hierarchy. Another series of finite size groups (usually r to simplify) is formed and the procedure is repeated until a single representative emerges in the k^{th} round of votes (see fig. 2.1). The recursion procedure can be mathematically described using binomial laws. Let p_i be the probability of an elected representative on level i to hold opinion A,

$$p_{i+1} = \begin{cases} \sum_{l=\frac{r+1}{2}}^{r} \binom{r}{l} p_i^l q^{r-l}, \text{ if } r \text{ is odd} \\ \frac{1}{2}\binom{r}{\frac{r}{2}} p_i^{\frac{r}{2}} q^{\frac{r}{2}} + \sum_{l=\frac{r}{2}+1}^{r} \binom{r}{l} p_i^l q^{r-l} \text{ if } r \text{ is even} \end{cases}$$

Conclusions and applications

One of Galam's conclusion [15], which he claims to be "counterintuitive" is the "empirical difficulty in changing leadership in well established institutions". Indeed, Galam's model features a bias towards established opinion in case of local ties when a cell of r even individuals have a $50:50$ ratio between A and B. In the case where B corresponds to the current political situation (eg. the United Kingdom is part of the European Union) and there is a local tie between A and B, B wins the local election as according to Galam [15]: "in case of no decision, things stay as they are". Although disputable, this assumption leads to the general result that in order to have a good chance to change leadership in a well established institution, the opposite opinion must have a substantial portion of supporters which is close to 77%. In most cases of opposition representing less than 77% of the population, the established opinion won the hierarchical election in very few rounds.

Using these results, Galam provided in 1997 [16, 17] the necessary conditions for the election of the French extreme right party Front National. His conditions were met in 2002 when the Front National leader Jean-Marie Le Pen won the first round of the French presidential election [18] before losing the second round [19].

2.2.5 United we Stand Divided we Fall (Sznajd)

Also based on the Ising model, USDF was described by Sznajd [20] in 2001 to model opinion dynamics in closed (isolated) communities. Is considered a population of agents which have to stand for either of two options A or B. The model studies the time evolution of m a parameter which describes the difference between agents in favour of A and those in favour of B. Agents are arranged in a Ising spins chains (one dimensional network), $(S_i; i = 1, ..., N)$ and obey the following rule:

(i) if $S_i S_{i+1} = +1$ that is two consecutive agents have the same opinion, both S_{i-1} and S_{i+2}adopt their opinion

(ii) if $S_i S_{i+1} = -1$, that is two consecutive agents have a diverging opinion, S_{i-1} takes S_{i+1}'s

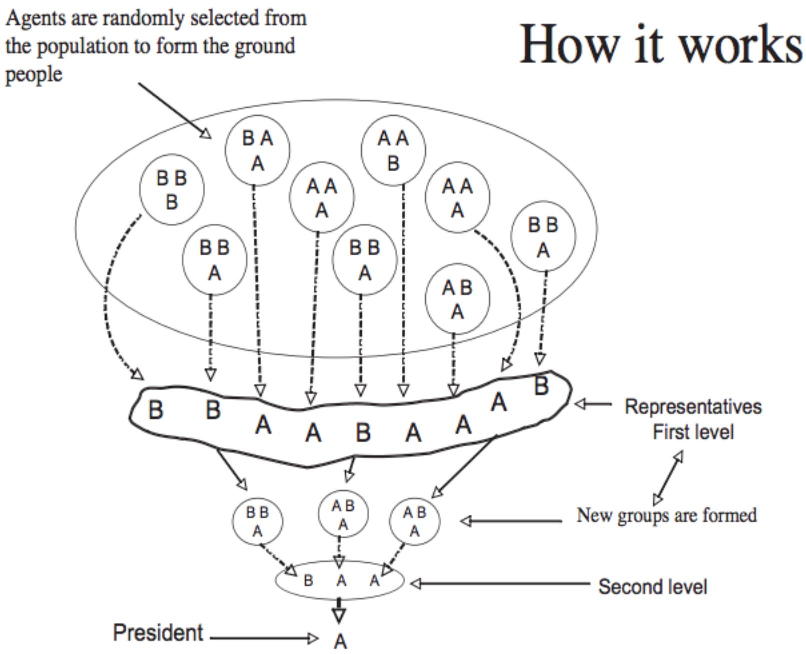


Figure 2.1: Three round voting hierarchy with group size $r = 3$, from *Sociophysics: a review of Galam models*" [15]

opinion and S_{i+2} adopts S_i's opinion
According to Sznajd, these rules "describe the influence of a given pair on the decision of its nearest neighbours". If two neighbouring agents share a common opinion, their respective neighbours will follow. On the other hand, if two neighbouring agents are divided, their neighbours will mirror this disagreement.

Conclusions

Using standard Monte Carlo Simulation, the following conclusions were reached by the authors. First, in an isolated (closed) community as the one in which the model is set, only two final states can be reached: global consensus (called "dictatorship") or stalemate. Either all agents adopt opinion A or B, or their opinion divides equally to a ration of $50 : 50$. Second, a small minority can under some circumstances bring to a stalemate in the final state. Furthermore, to have a good chance of winning (above 50%), a group must represent more than 70% of all individuals, supporting Galam's result that to obtain change, the will for change must be shared by a majority of agents [14].

2.2.6 Comments

The discrete models described in this section are mostly interesting in their extreme simplicity. Using binary variables to account for individual's opinion allows to model the situation using very simple statistical physics formulas as seen in strike model (Section 2.2.1) or probabilistic models with a binomial distribution as in the case of the hierarchical model (Section 2.2.4). Having few parameters allows to finely predict how tuning them affects the evolution of the opinion and the end results, but involves a trade-off in terms of realism. Indeed, the binary characteristic given to agents' opinion cannot account for any nuance in their position: agents are either extremely for or extremely against a decision, with no room for indecisiveness or hesitation. Furthermore, some models choices seem quite arbitrary. In Galam's hierarchy model (Section 2.2.4), according to how agents are selected to form groups, the end result can be completely different. A similar arbitrary decision can be observed in Sznajd model (Section 2.2.5): why, when two consecutive agents have a diverging opinion, would their respective neighbour take the opposed opinion rather than conform? Given our objective, which is to model accurately the evolution of opinion in a population in the light of an upcoming referendum, binary variables appear to be inappropriate to answer our needs. Individuals can hold a range of opinion which can go from an extreme refusal to an extreme support, while others can be indecisive or hold middle-range opinion (being pro or con but in a moderate way). Continuous opinion models, described in the next section, put forward some elements that will be of interest to meet our requirements.

2.3 Continuous opinion modelling

The models described in this section use continuous (bounded) variables to model each agent's opinion. Nuanced position can therefore be taken into account, as well as cognitive biases and inner-group influences. Models that arise can therefore be quite complex but closer to real-situations.

2.3.1 Deffuant model

The model

Deffuant et al. [21] took a step further in modelling opinion dynamics by taking into account a psychological component inherent to human agents which is confirmation bias (see section 2.1.1). Two individuals interacting will update their opinion and reach a compromise only if their respective opinions differ by less than a certain threshold. The model can be more rigorously framed as follows. We consider a population of N agents labeled i. Each agent has a continuous opinion x_i ranging between 0 and 1. At each time steps two agents in the population are chosen at random and meet. We define a threshold d, a convergence parameter $0 < \mu < 0.5$ and let x_j and x_k be the agents' opinion. If $\mid x_j - x_k \mid < d$, their opinion will update as follows:

$$\begin{cases} x_j = x_j + \mu(x_k - x_j) \\ x_k = x_k + \mu(x_j - x_k) \end{cases}$$

Initial conclusions

After running computer simulations and varying μ and d parameters, Deffuant et al. reached the following conclusions [21], assuming initially uniformly distributed opinions in the population. For high threshold $d = 0.5$ and constant μ population tends to converge towards a half-way consensus at $\frac{1}{2}$ (recall continuous opinion ranges between 0 and 1). Lower threshold $d = 0.2$ implies the formation of two distinct clusters at 0.25 and 0.75. Changing influenceability parameter μ changed how wide apart were the clusters when they did form. Formation of clusters can be explained as agents with close opinions interact and reach consensus as their opinions converge. The long-term result is exchange only being possible within clusters as agents from two different clusters will have too much difference in opinion (which does not occur with a high threshold d which covers too much opinions in the population). These results suggest a possible clue as to why certain social circles or spatial regions will tend to act as echo chambers for specific opinions.

2.3.2 DeGroot: individual influence in a group

In order to study the conditions under which a group might reach a consensus, Morris DeGroot introduced in 1974 a model which accounts for influence each individuals might excert on one another [22]. We consider a population on N individuals and represent for each agent i his initial opinion θ_i in a $n \times 1$ vector $\boldsymbol{\theta}$. In a real life scenario, each person, according to his professional or academic background, relationship with others or charisma has an influence p_{ij} on all other j individuals, where $j = 1, ..., N$ and p_{ii} corresponds to how much an agent's values his own opinion. We can then use an $N \times N$ row-stochastic matrix $\boldsymbol{P}$ where each entry P_{ij} is this influence p_{ij} under the rule that all $p_{ij} > 0$ and $\sum_j^N p_i = 1$ for all j (that is the matrix is row-stochastic).

Opinion evolution

Opinion formation at each time step t for a given agent i with opinion θ_i follows the rule:

$$\theta_i(t+1) = p_{i1}\theta_1(t) + p_{i2}\theta_2(t) + ... + p_{in}\theta_n(t)$$

Each agent adjusts his opinion at $t+1$ as a weighted average of weight a_{ij} for agent j with opinion $\theta_j(t)$ at time t.

Markov process

The evolution of the population's opinion $\boldsymbol{\theta_t}$ at any discrete time-step t can be easily computed as a Markov process:

$$\boldsymbol{\theta}(t+1) = \boldsymbol{P}\boldsymbol{\theta}(t) \tag{2.1}$$

or

$$\boldsymbol{\theta}(t+1) = \boldsymbol{P}^{t+1}\boldsymbol{\theta}$$

We can illustrate this process with the following example from Dr. Arceneaux' lecture at Princeton [23]. Let there be a population of three agents A, B and C represented by the weighted graph in figure 2.2. The corresponding matrix of influence is: $\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ We set at time 0 the following vector of initial opinion $\boldsymbol{\theta_0}$= $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

The network evolves as follows: $\boldsymbol{P}^2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \end{bmatrix}$, $\boldsymbol{P}^3 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$, ...,$\lim_{t \to \infty} \boldsymbol{P}^t = \begin{bmatrix} \frac{2}{5} & \frac{2}{5} & \frac{1}{5} \\ \frac{2}{5} & \frac{2}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{2}{5} & \frac{2}{5} \end{bmatrix}$

and so at $t \to \infty$ we obtain $\boldsymbol{\theta_t} = \boldsymbol{\theta_0}\boldsymbol{P}^t = \begin{bmatrix} 0.4 \\ 0.4 \\ 0.4 \end{bmatrix}$ and consensus was reached in the network.

Consensus conditions

DeGroot also established under which condition consensus would be reached in a particular weighted influence network [22]. We will briefly present these conditions without diving into the demonstrations. The condition was stated as follows: "If there exists a positive integer n such that every element in at least one column of the matrix $\boldsymbol{P}^n$ is positive, then a consensus is reached". That is,

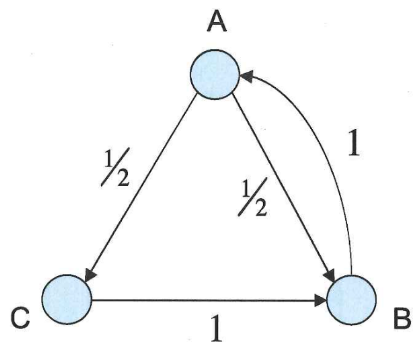


Figure 2.2: Example of a weighted network of individuals from *Introduction to Abstract Algebra with Applications to Social Systems* lecture by Dr. Arceneaux at Princeton

consensus is reached if the Markov chain is aperiodic and recurrent states communicate with one another.

2.3.3 Friedkin-Johnsen Model

Similarly to the DeGroot model, the Friedkin-Johnsen (FJ) models [24, 25] use a stochastic matrix to represent social influences $\boldsymbol{P}$ and a column vector $\boldsymbol{\theta}$ to represent the population's opinion. A diagonal matrix $\boldsymbol{\Lambda}$ =diag$(\lambda_1, \lambda_2, ..., \lambda_n)$ is introduced, where each $\lambda_i \in [0, 1]$ is the influenceability of each agent i, that is his willingness to update his own opinion when interacting with other agents. Traditionally, prejudice is assumed as $u = \boldsymbol{\theta}(0)$: the prejudice of each agent is his initial opinion before debates and interactions begin, possibly accounting for the agent's susceptibility to the anchoring bias (see section 2.1.2).

Opinion evolution

The population's opinion evolves at each discrete time step t according to the following rule:

$$\boldsymbol{\theta}(t+1) = \boldsymbol{\Lambda P \theta}(t) + (\boldsymbol{I} - \boldsymbol{\Lambda})u \tag{2.2}$$

where $\boldsymbol{I}$ is the identity matrix and u is a constant vector of agent's prejudices. In the case where $\boldsymbol{\Lambda} = \boldsymbol{I}$, 2.2 becomes

$$\boldsymbol{\theta}(t+1) = \boldsymbol{P\theta}(t)$$

the DeGroot equation (recall 2.1).

2.3.4 Comments

The first reason to find interest in continuous models is the use of continuous variables, as argued in Section 2.2.6. Second, some of these models as Deffuant (Section 2.3.1) and Friedkin-Johnsen (Section 2.3.1) take into account cognitive biases to which individuals are subject, adding realism to the model. They however remain extremely simplistic as they only take into account two or three parameters at most (stubbornness, original opinion anchored and mutual influence) which is far from all the factors that can influence an individual's decision and opinion formation. Far from being useless to answer our needs, these models will provide a first basis to construct our model.

In a first time, mimicking the Deffuant or FJ model will allow us to draw some conclusions before complexifying and moving away from this simple starting point.

2.4 Other literature of interest

We here describe a model of decision making which concerns animals (bees) rather than humans. Although it does not concern our target, it can still provide insight as to how deadlock resolution can be modelled.

2.4.1 Deadlock and consensus reaching: the case of bees

Opinion dynamics and modelling does not stop to human populations. Many researchers have successfully derived equations to study the behaviour of animal populations, such as fish schools [26], bird flocks [27] or insect swarms [28]. One case of interest is the one of house hunting honeybee swarms and the decision-making mechanism by which they choose the best potential nest site out of two locations. Once scouts have found a appropriate location for a potential new nest, they report back to the colony via a complex dance pattern [29] used to describe its characteristics and other bees deliberate to reach a decision as to which location will be favoured. The following model was defined and refined by Schleger *et al.* [30, 31], described in its simplified two alternatives form by Naomi Leonard during her presentation at the IEEE Conference on Decision and Control in Melbourne [32]:

$$\begin{cases} \frac{dy_A}{dt} = \frac{-y_A}{v_A} + v_A y_U + v_A y_U y_A - \sigma y_A y_B \\ \frac{dy_B}{dt} = \frac{-y_B}{v_B} + v_B y_U + v_B y_U y_B - \sigma y_A y_B \end{cases}$$

Where y_A, y_B are the proportion of scouts recruiting to respective locations A and B while $y_U = 1 - y_A - y_B$ describes the proportion of uncommitted scouts. v_A and v_B are two variables which describe the respective "quality" of each location. We are especially interested in σ, which we will describe later. Under normal circumstances, we have $v_A < v_B$ (or the opposite), that is one site presents more qualities than the other and should for that reason be favoured by the scouts, which is what Leonard verified [32]. The interesting case comes up when $v_A = v_B = v$, that is both sites are similar in quality and scouts cannot reach consensus in their decision process. The model then becomes:

$$\begin{cases} \frac{dy_A}{dt} = \frac{-y_A}{v} + v(1 - y_A - y_B)(1 + y_A) - \sigma y_A y_B \\ \frac{dy_B}{dt} = \frac{-y_B}{v} + v(1 - y_A - y_B)(1 + y_A) - \sigma y_A y_B \end{cases}$$

σ is the rate at which scouts emit a "stop-signal" which purpose is to shorten the deliberations and force the population to come up with a decision. Under weak σ, the population remains at indecision point while as σ increases, the population will shift towards one decision or the other, forcing a consensus. The threshold of the σ parameter above which consensus is reached was calculated by Schleger *et al.* [30, 31] to be:

$$\sigma^* = \frac{4v^3}{(v^2 - 1)^2}$$

An illustration of the phenomenon is shown in Figure 2.3 where we see how below threshold σ^* the scouts do not reach consensus and are in a deadlock, unable to reach a decision. Above the threshold, the opinion becomes unstable (red line) and the population tends to reach a consensus

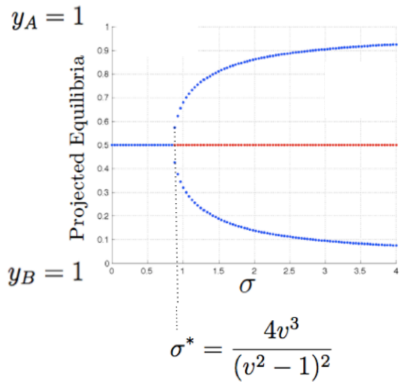


Figure 2.3: Pitchfork bifurcation of opinion, consensus reaching depending on stop-signal rate, from Naomi Leonard's lecture *Learning from Nature: When Models Meet Multitude* [33]

randomly on either of the two options.

The interest for this model lies in the variable σ which pressures the population into reaching a decision rather than waiting in a deadlock situation for a better opportunity to arise. Such a variable could be used to model a deadline approaching during a referendum or an election, but will not be used in this project.

2.5 Further literature (for the interested reader)

The following papers have not been included in the previous sections to keep the content clear and concise but are still worth mentioning as they provide complementary information and allow one to explore mentioned topics more in depth.

2.5.1 Further reading on discrete opinion modelling

The interested reader who wishes to learn more about Galam's work on modelling opnion can read his 2008 paper "*Sociophysics: a review of Galam models*" [15] in which he summarises 25 years of his own research on social dynamics which go beyond the scope of the considered models, proposing variations and including models for passive support to terrorism and formation of countries coalitions and fragmentations. Further information on statistical physics and its applications (especially in terms of opinion and crowd dynamics) can be found in the comprehensive literature review "*Statistical physics of social dynamics*" by Castellano et al [34].

2.5.2 Further reading on discrete continuous opinion modelling

A good introduction to the subject which helped me understand the topic and which curious reader should study is the literature review by Prokurnikov *et al.* [35] which also introduces all the related mathematical concepts which are involved in the domain. This paper will be used in writing the final report when setting up the mathematical part of the background section. Readings on applications of continuous opinion models include a project on Deffuant model in social networks [36] where "employed agents" have been introduced which act to shape opinion of individuals in the graph without updating their own.

2.5.3 Further reading on other topics

An excellent introduction to the topic of modelling animals behaviour can be found reading Lett and Mirabet's literature review on the subject [37]. The full description of honeybees nest-seeking behaviour can be found in Naomi Leonard's presentation at the IEEE Conference on Decision and Control in Melbourne [32] where she details the dance by which scouts communicate their findings.

Chapter 3

Analysis and Design

3.1 High level view

We here introduce a very simplistic and high level view of the model. This description is far from depicting the model's inner-workings - where most work and research was accomplished - and should only be used as a means to grasp the "big picture". All the relevant details will be found in Section 4.

We model a population composed of several sub-groups each containing the same amount of agents, whose opinions are distributed following certain distribution parameter at initialisation (See Sections 4.4 and 4.37).

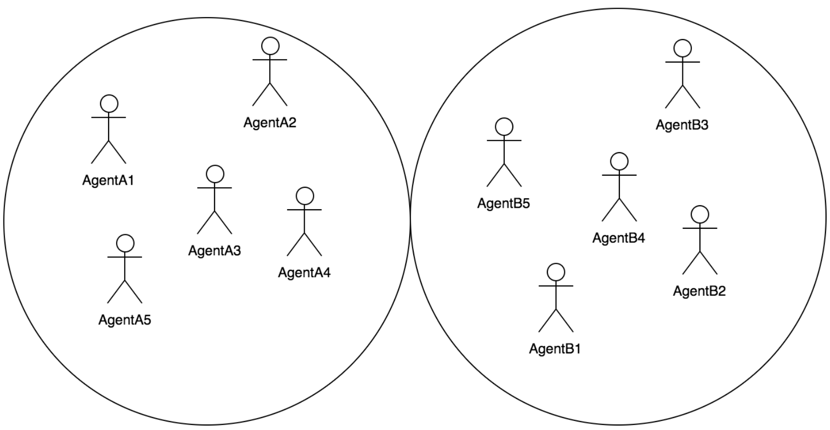


Figure 3.1: High level view - two sub-groups each containing five individuals

At each iteration, two agents from each group are selected to interact following a set of rules described in Section 4.

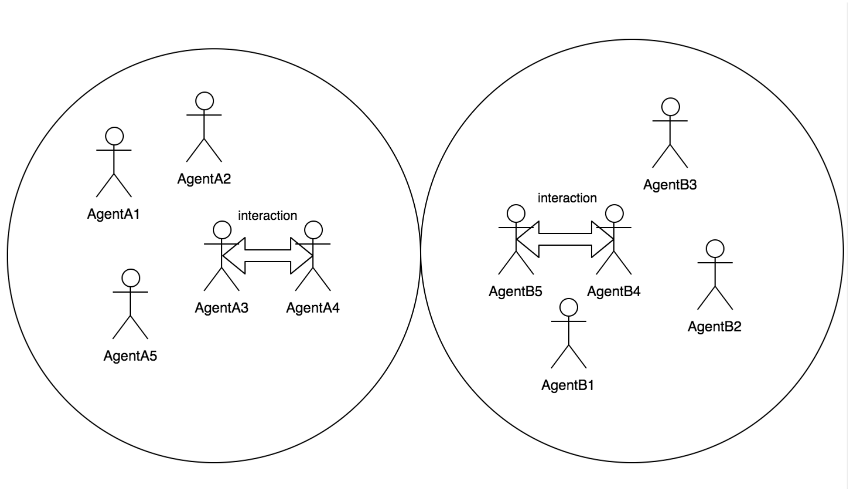


Figure 3.2: High level view - agents interacting inside sub-groups

Once interactions have been handled and opinion updated (if stubbornness permits, see Section 4.3), two subgroups are selected at random from the population from which two agents will be swapped to account for social mobility.

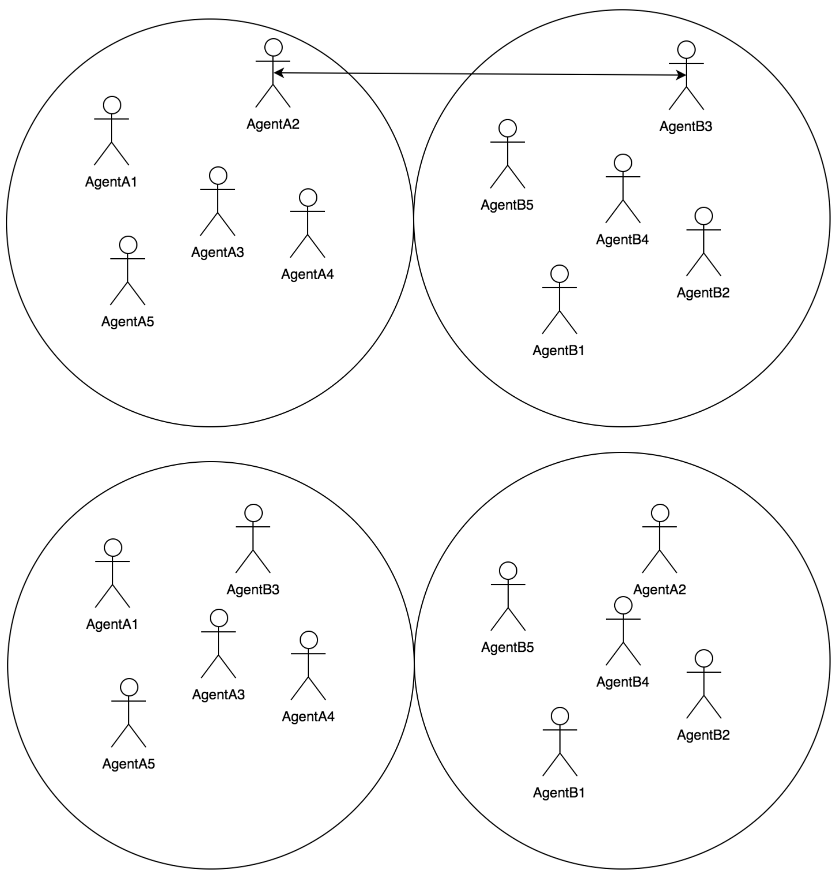


Figure 3.3: High level view - social mobility

3.2 Implementation

The model was implemented in *MATLAB* as it is one of the state-of-the-art software and programming language when it comes to modelling. The code for the final model is available on *GitHub* at the following address: https://github.com/victoralxdr/modellingBrexit/

3.2.1 Simple models - 3-dimensional array implementations

The first development stages of the model adopted an array-oriented approach to fully take advantage of *MATLAB*'s ability to compute large arrays quickly and efficiently.
The population was represented in a 3-dimensional array *population(groups, agents, features)* with the *features* dimension holding opinion, stubbornness and age values. This allowed for quick computation of interaction between agents, easy implementation of social mobility and results presentation as relevant features could be plotted immediately from the population matrix. This approach was kept until age and distribution bias were introduced (see Section 4.4) at which point it became preferable to adopt a new method.

3.2.2 Later models - object-oriented approach

As the model was becoming more complicated and more features were becoming agent-specific (changing stubbornness and convergence rate, socio-demographic features), the code was becoming less-readable when using the 3-dimensional array approach making it tedious to locate mistakes, code-redundancies and debug properly when mistakes occurred. It was therefore decided to change the implementation method and to adopt an object-oriented approach. Adopting an object-oriented approach implied abandoning *MATLAB*'s computing efficiency since each at each new simulation, the population matrix would have to be filled with the corresponding number of instantiations of individuals. Social mobility, which in the array-oriented approach consisted in copying a few integer values now implied copying and swapping objects which is again more computationally complex. Accessing and modifying the attributes of an object stored in a matrix demands more operations than the straightforward access of an integer value stored in a matrix involving again more complexity. It however permitted better code readability, more straightforward model upgrading and a more intuitive approach to the whole problem.
An "individual" class was created to encapsulate all agent-related attributes which were: age, bias, educationLevel, gender, income, mu (convergence rate), opinion and stubbornness (see below, figure 3.4). Specific methods were implemented to construct "individuals" objects and deal with Deffuant interactions when two "individuals" objects interact during the simulation.
Although it is not considered good practice and led to a a few attributes update mistakes, attributes were kept public and external functions were defined and called by the constructor to set their respective values. This allowed for an easier implementation of the three different evaluation models, since the "individual" class would remain unchanged as well as the main script, only changing the attributes distribution functions when needed.

3.3 Development protocol - incremental and iterative approach

The chosen protocol is a mix of an incremental and iterative approach, which consists in designing, implementing and testing the model incrementally (small steps by small steps) until the final model

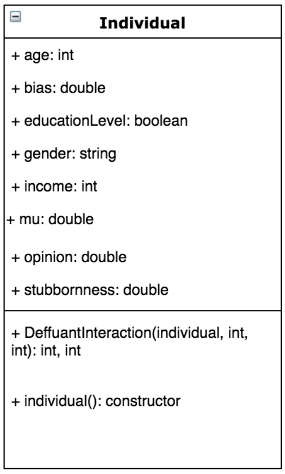


Figure 3.4: ULM class diagram of Individual (made with draw.io)

is achieved, while reviewing previous steps when necessary. We start by implementing a simplistic model with very few features, which will be seldom realistic but easy to implement and analyze in terms of results. The goal is then to incrementally add more features to the model, improving initial distribution of opinion, interactions between agents, willingness to update opinion etc. Each step includes experiments to study the benefit and relevance of each additions and possibly reveal flaws which will need to be corrected (eg. section 4.6). Along the way, some features were dropped because they did not behave as expected (see section 4.6.2).

3.4 Experimenting Protocol

Unless stated otherwise, for the development part, the model will always run according to the following guidelines. A population of $N = 1000$ agents is divided into g groups of $p = \frac{N}{g}$ individuals, either randomly or non-randomly. At initialization, each agent is given an opinion following a given distribution which will vary according to the model. The model will run for $t_{max} = 1000$ iterations, with initialization at t_0. At each iteration (with exception for model 1 4.1 two agents will be chosen at random from each group g_i and interact with each other. We will run the model for a given number of time which will be specified, usually $simulation \in [100, 1000]$ which should allow for rigorous statistical analysis of the results in terms of population opinion. Number of agents and iterations are kept at 1000 for the first milestones as it allows for fast running model and quick implementation and result analysis. Later models will be using larger population and more iterations to draw conclusive insight and results, more representative of an actual population.

3.5 Metrics

The following metrics were used to asses each model and improvements.

3.5.1 Statistical estimators

Statistical estimators are measured at t_0 and t_{max} and averaged on all simulations to get an insight of how model performs on a high number of runs.

Global population

Main parameters are *population mean* and *population standard deviation* which measure respectively the average opinion of individuals on the overall population of N agents as well as the standard deviation. Their purpose will be to track how the overall population's opinion evolves, and the mixity of opinions in the population - for example, mean and standard deviation can reveal if the population's opinion converges as in Section 4.1.

Group population

Similar measure is done at the end of each simulation measuring the average standard deviation of opinion within each group, averaged over all simulations. This will provide knowledge on opinion spread within each social group simulated and compare it to opinion distribution on the overall population, allowing to account for any opinion echo-chamber or instead inner-group distribution being similar to population opinion distribution.

3.5.2 Referendums

Some models will be assessed by running a "referendum" at t_0 and t_{max}. Each agent's opinion will be normalized as follows:

$$opinion = \begin{cases} -1, & \text{if } opinion < 0 \\ 1, & \text{if } opinion > 0 \end{cases}$$

The aim is to reflect how an agent would place his vote in a referendum according to his current opinion. This will allow to study more intuitively the population's opinion over iterations. It is however important to keep in mind that simulated referendums only indicate the proportion of the population situated below or above the 0 opinion line, and does not inform on overall opinion distribution, existence of opinion echo-chambers and extreme opinions.

Chapter 4

Analysis and design

4.1 Group mean opinion

Model design

The first proposed model is a continuous opinion model inspired from Galam's group majority model (see section 2.2.3). Unlike Galam whose model forms groups randomly at each iteration, this model forms the group at random at initialization and will keep them somewhat consistent to account for the fact that individuals tend to stick to their social group. Furthermore, for simplicity, each agent is assigned at t_0 an opinion following uniform distribution $\sim U(-1, 1)$.

Social interaction and opinion formation

In a very simplistic and somewhat unrealistic fashion, at each iteration t_i, agents in each group will form opinion following an extreme groupthink effect (see section 2.1.4) and adopt the group's average opinion.

Social mobility

The model accounts for the fact that individuals while tending to stick to their social groups can also be subject to social mobility, for various reasons such as changing jobs, moving from house, having a new occupation etc. A "swap" boolean variable is introduced. If swap = 1, at each iteration t_i two agents from two different groups are swapped to introduce some social mobility. Although it is now how social mobility takes place in real-life - people rarely exchange their place with someone else from another group - this is an easy way to implement social mobility given our array-oriented approach of the problem (see Section 3.2).

Running the simulation

The model is ran as described in Section **??**, first with swap = 0 and then swap = 1. We repeat the simulation 1000 times. The opinion of each agent is studied at times t_0 and at the end of the simulation at time t_{1000}.

Results for swap $= 0$

When swap = 0, that is groups stay as they were formed at initialization, the following results were obtained:

	t_0	t_{1000}
average sample mean	4.012×10^{-4}	4.012×10^{-4}
average sample standard deviation	0.5771	0.0791

Table 4.1: Population opinion at t_0 and t_{1000} over 100 simulations

The sample mean and standard deviation values of the population's initial opinion correspond to the theoretical values we obtain under the uniform distribution $U(-1, 1)$ that is $\mu = \frac{1+(-1)}{2} = 0$ and $\sigma = \sqrt{\frac{(1-(-1))^2}{12}} \approx 0.5773$. As agents update their opinion taking the average of their group's opinion, it is obvious that the population's mean opinion will not evolve regardless of the number of iterations on which the simulation is ran. Furthermore, the standard deviation will shrink as observed since all agents adopt a new opinion close to the theoretical mean. As no agents change group in this simulation, the same results would have been obtained after a single iteration with no evolution afterwards.

Results for swap $= 1$

When swap $= 1$, that is agents can now move across groups causing steady evolution in social groups composition, the following results were obtained:

	t_0	t_{1000}
average sample mean	4.840×10^{-4}	4.840×10^{-4}
average sample standard deviation	0.5777	0.0100

Table 4.2: Population opinion at t_0 and t_{1000} over 100 simulations

We observe a similar initial distribution as in the previous simulation. Once again, the mean value of opinions does not evolve as agents adopt the group's mean opinion at each iteration. We however notice that the standard deviation at t_{1000} has decreased from 0.0791 to 0.0100 as social mobility causes each group's mean opinion to converge slowly towards the theoretical mean opinion. Having agents move from a group to another in this model causes a uniformisation and convergence towards the mean of the whole population opinions.

Discussing model results

In regard to our goal, we can address the same criticism as to the models described in the literature review (section 2): this model is oversimplistic. While the structure of the population and the concept of introducing sub-groups to model social groups individuals can be part of is adding realism and complexity, the interaction-system to which agents are subject is unrealistic and leads to a general consensus at average opinion which would not happen in real-life. We therefore build our second model attempting to account for the fact that two agents, even within a same social group might not reach consensus in their social interaction and adopt a mid-range opinion at all time.

4.2 Group Deffuant interactions

Model design

Building upon the second model, the same structure is kept where a population of agents is divided into fixed-size groups and a boolean variable controls whether or not agents are subject to social mobility. The interactions between agents of a same social group are now made more complex as they will follow the rules of the Deffuant model (see section 2.3.1). At each iteration t_i, two agents within each group are chosen at random and will update their opinion according to the Deffuant rules.

Simulation

We keep the protocol parameters, distribute uniformly opinion at initialization $\sim U(-1,1)$ and run simulation 1000 times first with low confidence bound $d = 0.25$ and then high confidence bound $d = 0.75$. The opinion of each agent is studied at times t_0 and at the end of the simulation at time t_{1000}. The convergence parameter will keep value $\mu = 0.5$ as it only influences the speed of convergence rather than the opinion distribution. Swap variable is also turned off and on to study its effect on the model.

Results: $d = 0.25$, $swap = 0$

For a low confidence bound and no agent swap, the following results are observed.

	t_0	t_{1000}
average sample mean	-8.213×10^{-4}	-8.213×10^{-4}
average sample standard deviation	0.5772	0.5566

Table 4.3: Population opinion at t_0 and t_{1000} over 100 simulations

The mean of the population's opinion does not evolve with each iteration as Deffuant interaction causes a slow convergence of each agent's opinion toward its group mean opinion. This is confirmed by the slow decay in standard deviation which shows how agents slowly see their opinion converging towards mean opinion. Looking at how the opinion varies from a group to another we obtain the following results:

	mean	standard deviation
average inner-group standard deviation	0.5566	9.352×10^{-4}

Table 4.4: Inner-group opinion standard deviation over 100 simulations

As agents do not move from a group to another, each group on average presents a standard deviation which is close to that observed in the full population, with a distribution of standard deviation very similar from one group to another.

Results: $d = 0.25$, $swap = 1$

Introducing social mobility, the final standard deviation of the population's opinion is a little lower than for no social mobility, which allows us to state that introducing social mobility causes the whole population's opinion to converge towards the average opinion.

	t_0	t_{1000}
average sample mean	-4.672×10^{-4}	-4.672×10^{-4}
average sample standard deviation	0.5774	0.5457

Table 4.5: Population opinion at t_0 and t_{1000} over 100 simulations

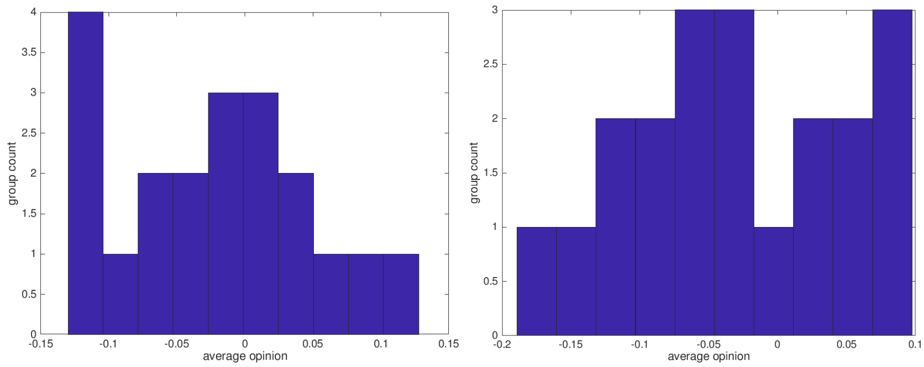
Looking at inner group standard deviation of opinion, similar results are obtained with respect to the case where the model has no social mobility.

	mean	standard deviation
average inner-group standard deviation	0.5451	0.0012

Table 4.6: Inner-group opinion standard deviation over 100 simulations

We can compare both models to see that the effect of social mobility is quite minor in this case:

Figure 4.1: Low bound $d = 0.25$

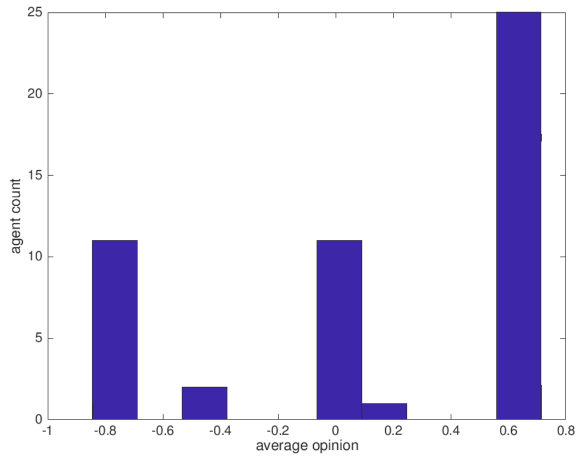


(a) Histogram after 1000 iterations, $swap = 0$

(b) Histogram after 1000 iterations, $swap = 1$

In both cases the opinion is quite spread among different groups.
"Diving" into a single group, we can however notice that under low confidence bound, agents form clusters of opinion which was to be expected considering results reached by Deffuant et al. when running simulations in a single group of agents (see section 2.3.1). A slight difference should be highlighted: under Deffuant's experiment, two clusters form a 0.25 and 0.75 when $d = 0.25$ while in our case four clusters form at -0.75, -0.4, 0.2 and 0.7 (close to the expected -0.75, -0.25, 0.25, 0.75) as our $d = 0.25$ and opinion ranges from -1 to 1 rather than 0 to 1.

Figure 4.2: Low bound $d = 0.25$ inner-group opinion distribution after 1000 iterations



Results: $d = 1.0$, $swap = 0$

Increasing the confidence bound to $d = 1$ has some significant effect on the population's opinion, even in the case where there is no social mobility, as the standard deviation of the population's opinion quickly reaches a value close to 0 as the agent's can interact with each other and see their opinion converging towards mean opinion.

	t_0	t_{1000}
average sample mean	3.925×10^{-4}	3.925×10^{-4}
average sample standard deviation	0.5774	0.0148

Table 4.7: Population opinion at t_0 and t_{1000} over 100 simulations

Looking at standard deviation inside each group, we can observe that on average it is close to 0 and we obtain the same results from one group to another. Increasing the confidence bound therefore causes opinion inside each group to collapse to its initial mean value.

	mean	standard deviation
average inner-group standard deviation	6.3×10^{-3}	1.778×10^{-4}

Table 4.8: Inner-group opinion standard deviation

Results: $d = 1.0$**,** $swap = 1$

Introducing social mobility does not change the final standard deviation of the opinion in the population.

	t_0	t_{1000}
average sample mean	-3.329×10^{-4}	-3.329×10^{-4}
average sample standard deviation	0.5773	0.0146

Table 4.9: Population opinion at t_0 and t_{1000} over 100 simulations

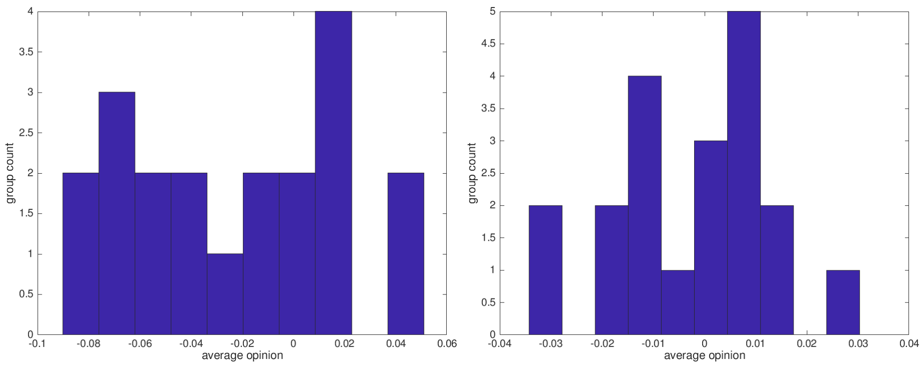
In terms of inner-group standard deviation, introducing social mobility does not change the distribution of inner-groups opinion.

	mean	standard deviation
average inner-group standard deviation	6.2×10^{-3}	1.568×10^{-4}

Table 4.10: Inner-group opinion standard deviation over 100 simulations

We can compare both models to see that the effect of social mobility is quite minor again:

Figure 4.3: Low bound $d = 0.25$



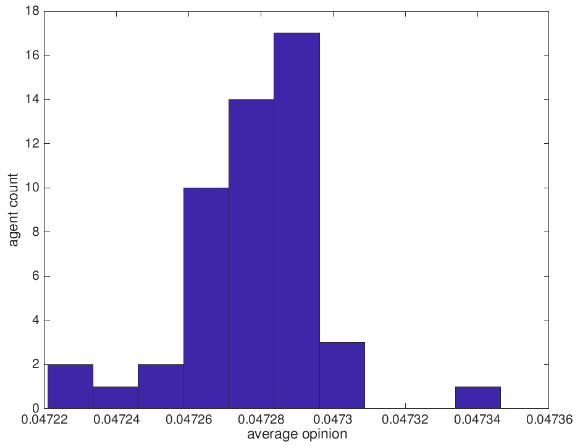
(a) Histogram after 1000 iterations, $swap = 0$

(b) Histogram after 1000 iterations, $swap = 1$

In both cases the opinion has converged and the groups hold opinion in the range $[-1, 1]$.

"Diving" into a single group, we notice that under high confidence bound, agents every agent has seen his opinion coverge strongly toward the overall mean, again as in the results reached by Deffuant et al. when running simulations in a single group of agents (see section 2.3.1).

Figure 4.4: High bound $d = 1.0$ inner-group opinion distribution after 1000 iterations



Discussing model results

Increasing the complexity of agents interactions did indeed lead us to more realistic results as in the case of low confidence bound, clusters of opinion could form within a social group which is more plausible when studying a real population. We now wish to take into account the fact that individuals may not always wish to update their opinion, introducing the concept of stubbornness.

4.3 Introducing stubbornness

Model design

This third version introduces the concept of stubbornness in agents: each agent is given a stubbornness characteristic according to which he will or not update his opinion on a Deffuant like model. Stubbornness of each agent is defined following the assumption that the more an individual holds an extreme position, the more stubborn he will be, while indecisive individuals can be more easily influenced into either direction depending on their encounter and social interactions. Stubbornness is assigned as follows:

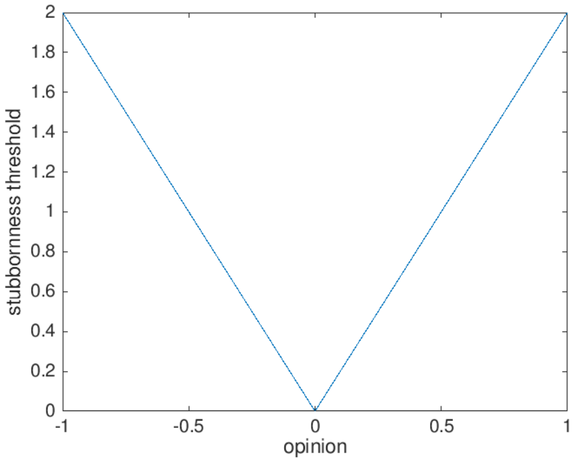
$$s = 2\times \mid opinion \mid$$

where *opinion* is the agent's current opinion. Under this stubbornness function, the more extreme is the agent's opinion, the less likely he is to update his opinion when encountering and interacting with a more moderate agent. ON the other hand, individuals with more moderate position can be more influenced by extreme individuals:

Simulation

We use again default protocol parameters and run the simulation 1000 times. At initialization opinion is distributed among agents following uniform distribution $U(-1, 1)$ and stubbornness is updated whenever an agent's opinion evolves. At each iteration two agents A and B from each group interact and update their opinion following the rule: If $\mid x_A - x_B \mid < s_A$ then $x_A =$

Figure 4.5: Initial stubbornness function



$x_A + \mu * (x_B - x_A)$ where x_A and x_B respectively denote the opinion of agents A and B and s_A is the stubbornness of agent A. Agent B's opinion is updated in a similar fashion but according to parameter s_B. We keep social mobility at each iteration.

Results

Following results were obtained:

	t_0	t_{1000}
average sample mean	-1.628×10^{-4}	0.0022
average sample standard deviation	0.5772	0.6111

Table 4.11: Population opinion at t_0 and t_{1000} over 100 simulations

Unlike previous models, overall mean opinion differs at iteration t_{max} from initial value, diverging slightly away from 0. This is a significant achievement as it confirms that our model population's opinion has evolved over time which appears to be more realistic than a population holding the same opinion over time on average. Furthermore, standard deviation has increased over iterations which also did not occur before (previous models saw standard deviation decreasing) as now the population's opinion seems to be gaining in diversity rather than converging to the mean opinion. This can be explained by stubbornness of agents which causes the most extreme agents to rally more naive and moderate agents to their cause. Agents are not converging together anymore: unilateral opinion updates can happen thus explaining the change in opinion mean value and increase in standard deviation.

We now look at standard deviation of opinion within the 20 sub-groups:

	mean	standard deviation
average inner-group standard deviation	0.4487	0.0047

Table 4.12: Inner-group opinion standard deviation over 100 simulations

It appears that all groups have a similar opinion distribution pattern as they all have closely the same standard deviation in their agents' opinion, that is every groups similarly holds opinion plurality among its agents.

Discussion and sources of improvements

Going further, it would be sound to improve the stubbornness variable for each agent as it was assigned in a quite simplistic fashion. Stubbornness should be made dependent on more factors than simply opinion: we could imagine a model in which stubbornness would depend on age, education level and information availability for more realistic results. Furthermore, Moussaid et al. [38] have shown that within a group of individuals, stubbornness tends to increase after each interaction as individual gain feedback and collect knowledge on the issue. This will be taken into account in further models.

4.4 Introducing age and initial distribution bias

Model design

We now wish to take into account agent's age and the influence it might have onto their opinion formation. Age distribution will follow the results of the 2011 Census which are the most precise data available regarding the UK population [39]. Furthermore, age will have an influence on agent's prior opinion before interaction which will follow the results given by Becker et al. [40]. The distribution of age and vote intention goes as follows:

age	population proportion	remain	leave	bias
18-24	11.9%	73%	27%	-0.23
25-34	17.0%	62%	38%	-0.12
35-44	17.8%	52%	48%	-0.02
45-54	17.6%	43%	57%	+0.07
55-64	14.9%	43%	57%	+0.07
65+	20.8%	40%	60%	+0.1

Table 4.13: Age distribution and corresponding voting intention

Simulation

Model is ran with default protocol parameters. At initialization, age is distributed among agents in discrete range $[18, 65]$ following table 4.13. A bias is then introduced according to the agent's age range, as detailed in the same table. This bias is then added to the uniform opinion distribution to shift the agent's opinion according to his age before normalizing to ensure that opinions are still in range $[-1, 1]$. Stubbornness is assigned as described previously and updated at each iteration. The simulation is ran 100 times as computation is becoming more demanding.

Results

	t_0	t_{1000}
average sample mean	-0.0086	0.0023
average sample standard deviation	0.5879	0.8001

Table 4.14: Population opinion at t_0 and t_{1000} over 100 simulations

Adding a bias at initial distribution obviously shifts average opinion from usual 0 value as the new distribution does not follow the previously used normal distribution $\sim U(-1, 1)$ but rather adds a bias to it. We furthermore notice that the divergence from initial value is still present as mean opinion increases simultaneously as standard deviation.

	mean	standard deviation
average inner-group standard deviation	0.7717	0.0059

Table 4.15: Inner-group opinion standard deviation over 100 simulations

In terms of sub-groups, each sub group seems to hold relatively high diversity of opinion as in average each group has a standard deviation of 0.7717. Furthermore, this opinion diversity can be observed in every group as the standard deviation across all 20 groups is close to 0.
Now that opinion bias has been introduced at initialization and that opinion of the population is evolving rather than simply converging towards mean, studying results in terms of a "referendum situation" seems suitable for interpreting how well model is evolving. At each simulation, both at iteration t_0 and t_{max} a "referendum" takes place where individual's whose opinion is negative is counted as -1 (remain) while those with positive opinion are counted as $+1$ (leave). The following results are obtained over the 100 simulations:

	$remain$	$leave$
t_0	50.725%	49.275%
t_{max}	49.57%	50.43%

Table 4.16: Average referendum results over 100 simulations

The referendum results confirm what the mean and standard deviation measures suggested: opinion instead of converging towards a mean is instead shifting towards the "leave" position. The opinion gap is increasing over each iteration while the mean increases, which is likely due to extreme agents being stubborn, drawing towards them more moderated agents. Furthermore the drift towards the "leave" or $+1$, while the initial bias introduced with age distribution at initialization was tending toward "remain" or -1 (the initial mean being ~ -0.0086) confirms that the population's opinion is evolving.

Discussion

It is important to highlight that bias according to age assignment is only based on referendum outcomes rather than individual's opinion beforehand. That implies that the similarity between our model's results and the real-case scenario of Brexit referendum (recall result was 48.1 : 51.9 in favour of leave [2]) is likely to be due to a good opinion distribution at initialization which fits the end result of the referendum rather than the model's capability to forecast opinion evolution. However, it should be noted that the model is now able to reproduce a population's opinion at a given time with relatively close precision, meaning we can now focus on improving factors which influence opinion evolution.
Looking at diversity within each group which is relatively high (inner-group standard deviation mean $\sim$ 0.77, table.4.15) and remarking it is similar across all 20 groups suggests that group formation might have to be re-worked. Attention should be dedicated to less-random social group formation as individuals in real-life tend to bond with people who resemble them in terms of age, social background and sometimes political opinions.

4.5 Normal biased initial opinion distribution

Model design

Out of curiosity, uniform distribution was replaced by normal distribution at initialization, from the idea that out of a large population of individuals, following the central limit theorem, the distribution of opinion from a given age group should follow a normal distribution around its mean.

Simulation

Each age group is therefore assumed to have opinion distributed following $\sim N(\mu + bias, \sigma)$. We set $\mu = 0$ and $\sigma = \frac{1}{3}$ as by definition 99.72% of values are between $[-3\sigma, 3\sigma]$ which in this case gives $[-1, 1]$ ensuring that 99.72% of initial values will respect our opinion bound. We then shift each distribution by a bias computed according to the % of individuals of each aged group who voted for a particular option, as in the previous model. The simulation is ran 100 times for computational complexity reasons.

Results

	t_0	t_{1000}
average sample mean	-0.0081	-0.0437
average sample standard deviation	0.3496	0.1952

Table 4.17: Population opinion at t_0 and t_{1000} over 100 simulations

Adding a bias at initial distribution obviously shifts average opinion from usual 0 value while it still remains close to it. Initial standard deviation is $\sim$ 0.35 which is close to the standard deviation $\sigma = \frac{1}{3}$ used when distributed opinion among age groups. We however notice that while mean opinion diverges away from 0, standard deviations is decreasing by almost 50%.

	mean	standard deviation
average inner-group standard deviation	0.1432	0.0050

Table 4.18: Inner-group opinion standard deviation over 100 simulations

In terms of sub-groups, each sub group holds relatively lower diversity of opinion than in the previous models as in average each group has a standard deviation of 0.14. Furthermore, this opinion diversity can be observed in every group as the standard deviation across all 20 groups is close to 0.

The process of "referendum" is repeated.

	remain	leave
t_0	50.634%	49.366%
t_{max}	63.717%	36.2830%

Table 4.19: Average referendum results over 100 simulations

The referendum results shows that opinion instead of converging towards a mean is instead shifting towards the "remain" position, due to the combination of initial mean being less than 0 and stubbornness of agents which creates momentum in the direction of initial shift. We however obtain a much more polarized results than in the case of biased uniform initial distribution, as now a clear majority of the population has a negative opinion (< 0, voting for remain).

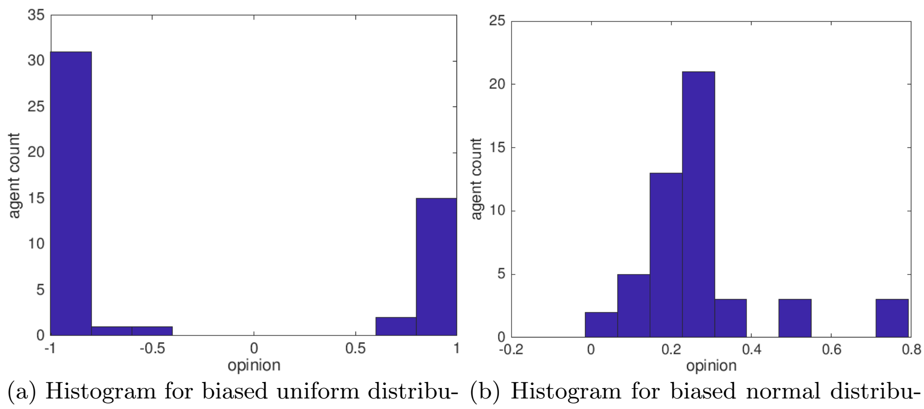
Discussion

As in the previous model, opinion and bias was distributed following referendum outcomes rather than individual's opinion beforehand. Once again, similarity between our model's initial results at t_0 and the real-case scenario of Brexit referendum (recall result was 48.1 : 51.9 in favour of leave [2]) is likely to be due to a good opinion distribution at initialization which fits the end result of the referendum rather than model's capability of forecasting opinion evolution. In the case of normal distribution, the end result is however far from modelling reality as one of the two sides has a lot more supporters, leading to a highly unbalanced ratio of 36 : 64.
Looking at diversity within each group which is relatively low (~ 0.14) and remarking it is similar across all 20 groups suggests that group formation might be similar to what can be witnessed in real life, where social groups tend to act as echo chambers for certain opinion. A trade-off between the two distributions appears. While uniform distribution yields overall closer results to reality with a balanced ratio, its inner-group dynamics show too much opinion diversity. On the other hand, normal distribution yields unbalanced results in the global population but allows for the formation of echo chambers in each subgroup.

Comparing uniform and normal distribution

We can support our statement looking at two subgroups in both cases of biased uniform distribution and biased normal distribution.

Figure 4.6: Agents' opinions within two groups after 1000 iterations



(a) Histogram for biased uniform distribution

(b) Histogram for biased normal distribution

The first group (fig. 4.6a) seems unrealistic since it is split into two clusters of highly opposed opinions. The second group (fig. 4.6b) shows some kind of consensus between individuals who hold similar opinion, with a few outliers who could be too stubborn to follow the others, demonstrating a typical echo chamber situation. Although we will prefer normal distribution, the ideal result would be a tradeoff between the inner group distribution of the normal distribution with the overall opinion obtained under uniform distribution. This statement makes the assumption that a realistic model will include social groups subject to groupthink featuring few outliers (see Section 2.1.4) while the overall opinion follows the EU Referendum results.

4.6 Re-calibrating parameters to obtain realistic results under normal distribution

Model design

Given the results obtained in section 4.5, normal biased distribution appears to be a promising initial distribution as it provides at t_{max} typical opinion distribution we expect within social groups. We however noticed a couple of weaknesses: the overall opinion drifts too quickly towards an extreme (it went from a ration of $\sim 49:51$ to $\sim 36:64$) and groups may be too large (although this is not specific to this distribution but shared between all models). We will therefore attempt to fit stubbornness parameters to reach more realistic results under normal biased initial distribution. In this new model, we will now study $g = 50$ groups of $p = \frac{N}{g} = \frac{1000}{50} = 20$ individuals (which seems more realistic in terms of social group size) and change the stubbornness function

$$s_1 = 2\times \mid opinion \mid$$

trying the following candidate functions to replace it:

function s_2	$2 \times \mid \sqrt{opinion} \mid$
function s_3	$2.63 \times \mid tanh(opinion) \mid$
function s_4	$2.57 \times \mid arctan(opinion) \mid$

Simulation

Each age group has opinion distributed following $\sim N(bias, 0.3)$ where bias is described in table 4.13. We first run our simulation for $g = 50$ and $p = 20$ to estimate model impact before modifying stubbornness function. We will then try all candidate stubbornness functions. The model will be ran over $t_{max} = 1000$ iterations and simulation will be ran 100 times.

4.6.1 Changing group size: results

Looking at the mean and standard deviation, we obtain slightly different results to the model with larger groups (50 individuals per group). While final means are almost equal for both group configurations, standard deviation increases when groups contain less individuals .

	t_0	t_{1000}
average sample mean	-0.0074	-0.0458
average sample standard deviation	0.3489	0.2663

Table 4.20: Population opinion over 100 simulations

Looking at opinion distribution within each groups, we can see that standard deviation within groups is on average smaller than when modelling the population with larger groups of individuals.

	mean	standard deviation
average inner-group standard deviation	0.0937	0.0081

Table 4.21: Inner-group opinion standard deviation over 100 simulations

We finally look at the consequence if running a referendum of opinion on this population, and notice that this time opinion has reached a slightly more balanced ratio.

	remain	leave
t_0	50.74%	49.26%
t_{max}	60.35%	39.65%

Table 4.22: Average referendum results over 100 simulations

Discussion - group size

In the light of the results obtained when varying group-size, it appears that a population with more small-sized groups sees its global opinion vary slower than when less larger-sized groups are present.

4.6.2 Candidate stubbornness functions: results

The original function is plotted in figure 4.7.

Running our simulation for the original stubbornness function for $t_{max} = 1000$ iterations, $g = 50$ groups of $p = 20$ agents, over 100 simulations, the following results were obtained:

	t_0	t_{1000}
average sample mean	-0.0074	0.0458
average sample standard deviation	0.3489	0.2663
remain	50.74%	60.35%
leave	49.26%	39.65%

Table 4.23: Summary of metrics over 100 simulations

	mean	standard deviation
average inner-group standard deviation	0.0937	0.0081

Table 4.24: Inner-group opinion standard deviation over 100 simulations

This results will be used to compare how suitable candidates functions are for establishing individuals' stubbornness.

Candidate 1 - $s_2 = 2\times \mid \sqrt{opinion} \mid$

For the first candidate stubbornness function, plotted in figure 4.8, we obtain the following results:

	t_0	t_{1000}
average sample mean	-0.0073	-0.1178
average sample standard deviation	0.3494	0.4645
remain	50.652%	61.805%
leave	49.348%	38.1950%

Table 4.25: Metrics over 100 simulations

	mean	standard deviation
average inner-group standard deviation	0.3166	0.015

Table 4.26: Inner-group opinion standard deviation over 100 simulations

Candidate 2 - $s_3 = 2.63\times \mid tanh(opinion) \mid$

For the second candidate function we get:

	t_0	t_{1000}
average sample mean	-0.0085	-0.0758
average sample standard deviation	0.3503	0.4132
remain	50.83%	58.959%
leave	49.17%	41.041%

Table 4.27: Metrics over 100 simulations

	mean	standard deviation
average inner-group standard deviation	0.2305	0.0147

Table 4.28: Inner-group opinion standard deviation over 100 simulations

Candidate 3 - $s_4 = \ 2.57\times \mid arctan(opinion) \mid$

The last candidate function yields following results:

	t_0	t_{1000}
average sample mean	-0.0085	-0.0758
average sample standard deviation	0.3503	0.4036
remain	50.83%	58.528%
leave	49.17%	41.4720%

Table 4.29: Metrics over 100 simulations

	mean	standard deviation
average inner-group standard deviation	0.2182	0.0138

Table 4.30: Inner-group opinion standard deviation over 100 simulations

Discussion: which candidate function is the best?

Looking at the original function and its three candidate replacements, we obtain some variability in the final results but no function seems to be superior to the others. Function s_2 behaves similarly to function s_1 but with much higher inner-group standard deviation, which was the reason for discarding uniform distribution. We therefore will not be using function s_2 for stubbornness. Functions s_3 and s_4 both behave similarly, with referendum results slightly more balanced than function s_1 but higher inner-group standard deviation of opinion. They furthermore lead to final standard deviation to be higher in the population. As none of the candidate functions seem to provide with a more realistic final distribution of opinion, we will be keeping function s_1 as our stubbornness function.

4.7 Convergence rate - decrease over time

Model design

We now seek to account for stubbornness building over-time as mentioned in section 4.3 following Moussaid et al.'s work [38]. We however explore a different method to do so: instead of having a

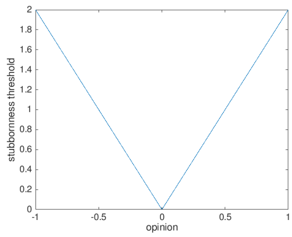


Figure 4.7: $s_1 = 2 \times \mid opinion \mid$

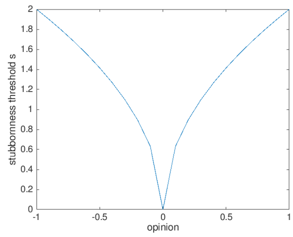


Figure 4.8: $s_2 = 2 \times \mid \sqrt{opinion} \mid$

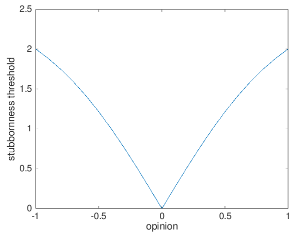


Figure 4.9: $s_3 = 2.63 \times \mid tanh(opinion) \mid$

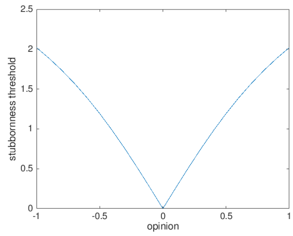


Figure 4.10: $s_4 = 2.57 \times \mid arctan(opinion) \mid$

gradually increasing stubbornness for each agent, we rather decrease the convergence rate μ (see section 2.3.1). This allows to stick to our assumption that stubbornness of an agent is linked to opinion, and account for the fact that over interactions, agents tend to stick with their formed opinion.

Simulation

We run over 100 simulations the model obtained in the previous simulation, keeping original stubbornness function with newly diminished group size. We incorporate the new convergence parameter as follows:

$$\mu_i = \mu_i \times (1 - t_i/t_{max}),\ i = t_0,...,t_{max}$$

Convergence parameter for the population therefore linearly decreases over each iteration. We therefore expect change in population opinion to decrease over time.

Results

	t_0	t_{1000}
average sample mean	-0.0083	-0.0142
average sample standard deviation	0.3498	0.2221
remain	50.83%	53.254%
leave	49.17%	46.746%

Table 4.31: Metrics over 100 simulations

	mean	standard deviation
average inner-group standard deviation	0.2105	0.0075

Table 4.32: Inner-group opinion standard deviation over 100 simulations

Discussion

This first model of decreasing convergence rate is simplistic in that all agents share the same convergence rate which decreases over time similarly for each agent. The following model will carry on our objective of realism by providing each agent with its own convergence rate μ which will evolve according to how many interactions each single agent has, instead of having it be common to all agents.

4.8 Convergence rate - agent assigned

Model design

We now expand upon work in section 4.7 by making the convergence parameter specific to each agent in the population. Since it was shown that agents tend to gain confidence in their own opinion the more they interact, it makes more sense to have an agent specific convergence rate which will diminish whenever an agent interacts than having a global parameter that will affect each agents the same way.

Simulation

We repeat section 4.7 experiment as described. The convergence parameter μ_i will be updated only when an agent a_i interacts (successfully or not, under stubbornness condition) following the rule:

$$\mu_i + 1 = \mu_i \times d,\ d \text{ decrease parameter} < 1$$

We will study the result under various values of d, expecting that as d decreases the final opinion distribution should be closer and closer to its original distribution.

Results

	$d = 0.75$	$d = 0.5$	$d = 0.25$
average sample mean	-0.0730	-0.0111	-0.0088
average sample standard deviation	0.2882	0.2599	0.2942
remain	63.482%	51.849%	51.413%
leave	36.518%	48.1510%	48.5870%

Table 4.33: Metrics over 100 simulations for all three values of d at t_{1000}

Discussion

All three parameters behave as expected: the smaller the value for d, the less opinion seems to evolve in the population. Ultimately, adding decreasing convergence parameter is equivalent to running the simulation over less iterations. For now, we will keep value $d = 0.75$ as it corresponds to how we would expect individuals to update their opinion gaining confidence while this value is not too small which would prevent the model from updating at all.

4.9 Adding to initial opinion distribution: gender, household income, education and political attention

Model design

Now that major parameters have been tuned - opinion dependent stubbornness, decreasing convergence rate, normal biased opinion distribution - we wish to add on factors influencing on agents initial opinion. Age was a good starting point, but cannot on its own for agent's opinion bias regarding a topic like Brexit. Studies and polls have shown that more factors were at play: education [40–43], gender [40, 42, 43], household income [40, 42–44] and political awareness [40, 42].

Initial design error and parameter discussion

The first attempt at adding these parameters to the initial opinion distribution was based on the erroneous assumption that the YouGov survey results [42] provided accurate measures of income, gender and political awareness distribution in the UK population. The goal of the survey was to provide insight as to how members of each of these categories (income, political awareness and gender) voted during the referendum, meaning that the size of each categories was not accurately representative of that same category in the UK population.
An example of this concerns education-related vote. While 26% of the survey interviewees had obtained a higher education degree, the actual statistic in the UK population is around 34% [45]. The model was therefore parametrized using wrong distributions and yielded erroneous results, which were noticed with the help of the statistical metrics used during experimentation.
It followed a re-consideration of the parameters and "political awareness", measuring how politically conscious agents are was dropped as it was difficult to find reliable sources of statistics measuring this in the UK population. Reliable sources were found for household income, higher education and gender through Official National Statistics [46] and a Guardian Review of higher education level [45].

Simulation

As when introducing age to the model (section 4.4), parameters are distributed among agents at initialization following tables 4.34, 4.35 and 4.36 adding biases for age, education, household income and gender together to shift the initial normal distributed opinion. Simulation is then ran following the same parameters as in the previous model (section 4.7).

education level	population proportion	remain	leave	bias
higher education achieved	34%	68%	32%	-0.18
higher education not achieved	66%	41%	59%	+0.09

Table 4.34: Education level and corresponding voting intention [42, 45]

household income (£)	population proportion (rounded to unit)	remain	leave	bias
< 21348	20%	38%	62%	+0.12
$\in [21348, 43190[$	40%	47%	53%	+0.03
$\in [43190, 64606[$	20%	58%	42%	-0.08
$\in > 64606$	20%	65%	35%	-0.15

Table 4.35: Household income distribution and corresponding voting intention [42, 46]

gender	population proportion	remain	leave	bias
male	50%	47%	53%	+0.03
female	50%	49%	51%	+0.01

Table 4.36: Gender and corresponding voting intention [39, 42]

Results

	t_0	t_{1000}
average sample mean	0.0054	0.0012
average sample standard deviation	0.3813	0.2513
remain	49.034%	49.301%
leave	50.9660%	50.6990%

Table 4.37: Metrics over 100 simulations

	mean	standard deviation
average inner-group standard deviation	0.2386	0.0071

Table 4.38: Inner-group opinion standard deviation over 100 simulations

4.9.1 Discussion

The obtained results provide with a realistic inner-group distribution while the overall population opinion is very close to the real-life Brexit scenario referendum, which is the ideal tradeoff we aimed for when comparing uniform and normal biased distributions (see section 4.5). Having reached this important milestone, we can now start considering implementations such as indecisiveness, turnout and impact of events.

4.10 Indecisiveness and turnout

Model design

While turnout for the referendum was 72.2% [2], meaning that 27.8% of agents did not vote, most surveys and polls also accounted for "indecisive voters" [47–49] which typically range from 3% to 15%. Several research papers debate the effect that turnout had in the referendum, and do not agree on which side it benefited [50, 51] or if it had any impact at all [40]. We however consider this feature as is would add comprehensiveness to the model.

Simulations

We start by incorporating indecisiveness in our model. Indecisive individuals are those who do not know conclusively where their opinion lies, which in the case of this model corresponds to neutral opinion, close to 0, on the spectrum of opinion we defined as ranging from -1 to +1. We arbitrarily define an indecisiveness threshold:

$$i = 0.05;$$

Individuals with opinion such that

$$\mid opinion \mid < i$$

will be defined as indecisive. It will be assumed that indecisive agents who vote on the day of the referendum will place a vote in a random fashion, which in the context of the model will be implemented at the end of each simulation, at t_{max} when measuring normalized opinion as a referendum. Every agent with "indecisive" opinion will randomly vote for -1 or 1, regardless of his final opinion at t_{max}. We parametrize the model as in the previous model (see section 4.9).
Regarding turnout, the model will simply randomly select 72.2% of agents' opinions when modelling the referendum at the end of each simulation.
We finally run a simulation where both indecisiveness and turnout are incorporated in the model.

Results

	remain	leave
EU referendum results [2]	48.11%	51.89%
result from previous model	49.301%	50.6990%
indecisiveness $i = 0.05$	48.88%	51.12%
turnout 72.2%	48.52%	51.48%
indecisiveness AND turnout	48.60%	51.39%

Table 4.39: Metrics over 100 simulations for all three values of d at t_{1000}

We only consider referendum at t_{max} as implementing indecisiveness and turnout only affects this final measurement of and does play a role in opinion formation and dynamics. Observing the results, we can see whether we implement indecisiveness, turnout or both bring the final referendum result closer from the real-life scenario result and increases the opinion-polarization. Turnout brings the strongest result, indecisiveness the lest significant and both result in a trade-off. The overall impact is still quite low as can be observed when plotting the distribution result (see fig. 4.11).

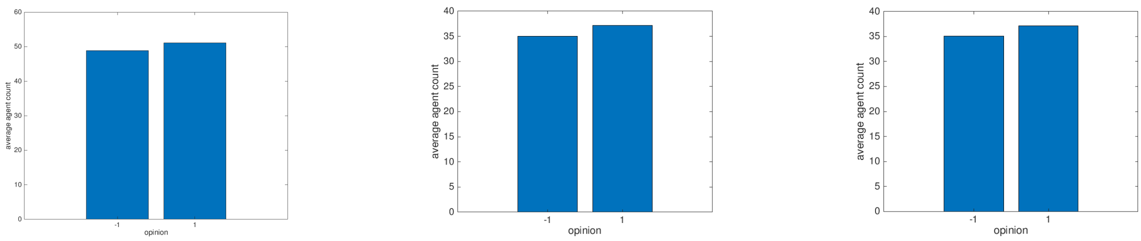


Figure 4.11: Comparing referendum results averaged over 100 simulations for the three models at t_{1000}: indecisiveness, turnout and combined (turnout changes the scale as 27.8% of agents were removed)

A possible improvement which will not be undertaken given remaining time is to study the correlation between indecisiveness and turnout in the scenario of a real referendum, as both parameters could be linked which would allow for a more realistic approach of the situation instead of assuming both to be uncorrelated as done in our model. It could for example be interesting to see if indecisive voters are less likely to vote than individuals with a stronger opinion.
These results allow us to conclude that our model now approaches closely the EU Referendum result we wished to model. Given the approach taken so far, the safest conclusion we can draw is that our model fits the end result after $t_{max} = 1000$ iterations on an average of 100 simulations with a low population of $N = 1000$ agents.

4.11 Events

Motivations and model design

We now introduce the last feature of the model, which is the impact of unusual, exceptional or shocking events can contribute in shaping a population's opinion. The first example that comes to mind when thinking of the EU Referendum is the murder of MP Jo Cox by a far right extremist on June 16, a week before the referendum [52, 53]. It has been shown that this particular event caused substantial emotional reaction on social media [54] and such emotion could be reflective of population's general opinion on related topics such as Brexit [55]. It is however still difficult today to asses which side between Leave or Remain "benefited" from this tragic event.
We will therefore look at two events which took place across the Channel: the terror attacks in Paris in January and November 2015. Both these events profoundly shocked the French public opinion. What is of interest for us is how both these events were correlated with a substantial rise of public approval of French President François Hollande [56–58] who on both occasion saw his national approval rate increase by 10 to 20 points (according to polls). These two changes in opinion are furthermore important as they both occurred after a strong political event, provoking a similar change in trend, when the President was quite unpopular (see fig. 4.12. We however wish to stress that the approval quickly decreased after the first event (Charlie Hebdo) but remained higher than its initial state, while after the second event it fell quickly and even decreased below the level before Charlie Hebdo.
Our model will attempt to account for such event by adding after a certain number of iterations t_{event} a push in opinion, shifting the opinion of a certain number of agents in a particular direction. Three different situations will be assessed: in the first one, all agents will see their opinion shifted by the same amount, in the second, the shift will be specific to each agent, accounting for the fact that in a population not every individual reacts in the same way to such events and in the third multiple events with various agent-specific shifts will occur in a single simulation.

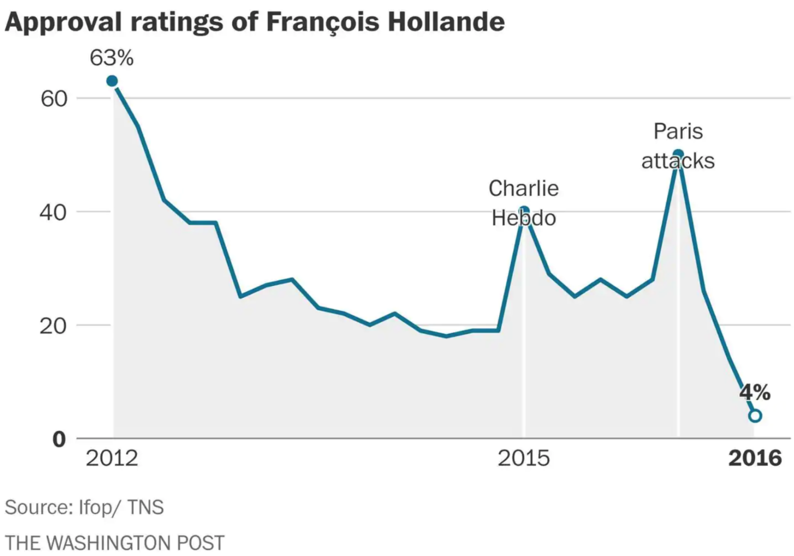


Figure 4.12: Francois Hollande's public approval [58]

4.11.1 Simulation 1: common shift

The first simulation will incorporate in the model an event triggered at mid simulation with gravity $gravity = 0.05$. Gravity is a bias which will be added to all agents' opinion in the population, in order to mimic an event which will shift the population's opinion a specific direction. A single simulation will be ran and opinion evolution will be displayed as a graph to allow for better visualisation of event-impact on opinion.

The described protocol did not yield the expected results as a single realization did not allow for averaging many simulations and providing statistically robust results which could be interpreted. It was therefore decided to trade computational efficiency for robustness and increase the population of agents from $N = 1000$ to $N = 10000$ with now $g = 500$ groups of $p = 20$ agents. These changes were first planned for when evaluating the model but have been applied at this stage of the development process to provide more robust results. This should have been done earlier in the process (probably from the very first model implementation, see discussion in conclusion, section 6).

First results and changing convergence rate

After running the first simulation, the following results came out: clearly showing that apart

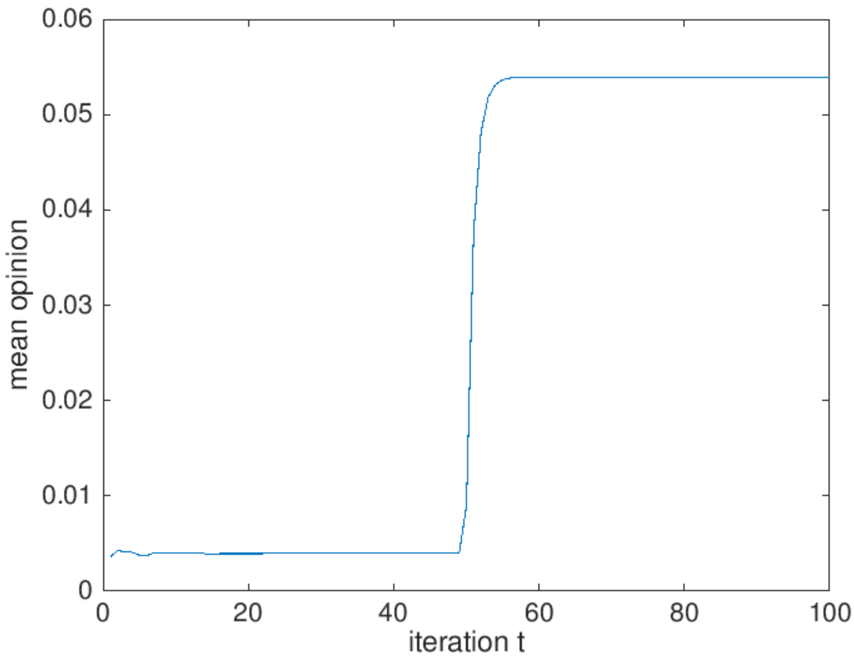


Figure 4.13: Mean opinion evolution when incorporating event with gravity 0.05

from the event at iteration 50, opinion stopped updating before iteration 10, which was quite unexpected. Careful investigation led to the following conclusion: a mistake in the code updated twice convergence rate for each agent, and convergence rate update was too strong and therefore lead to stable opinion in less than 10 iterations. If measuring mean opinion evolution over iterations had been done when setting convergence rate update, this mistake could have been avoided. This should have been expected in any case: considering 2 agents per group of 20 agents interact at each iteration, on 1000 iterations we expect each agent to interact on average $\frac{2}{20} \times 1000 = 100$ times which given $\mu_{i+1} = \mu_i \times 0.75$ gives a convergence rate of $0.5 * 0.75^2 \approx 0$ as $\mu_0 = 0.5$ which confirms that opinion would stop updating before the model stopped running.

Due to time constraints, experiments on updating convergence rate μ were done briefly and a new update rate was reached: $\mu_i + 1 = \mu_i \times 0.99$ to keep this feature while minimizing its negative effect as now $0.5 * 0.99^{100} \approx 0.2$. Using the new convergence rate update rule and event occurrence we get the following results:

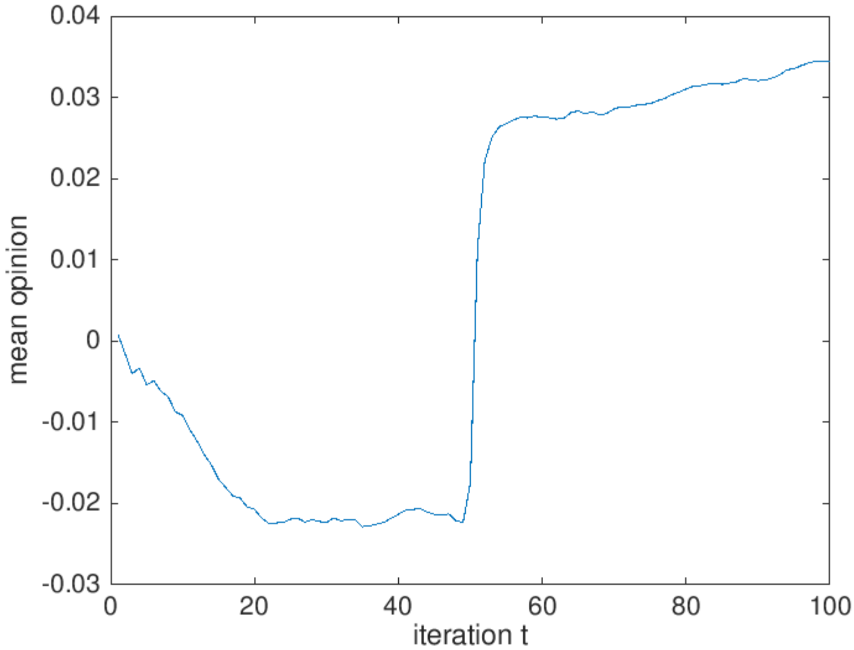


Figure 4.14: Mean opinion evolution when incorporating event with gravity 0.05 and new convergence rate rule

4.11.2 Simulation 2: agent-specific shift

The second simulation makes event opinion bias specific to each agent: when the event occurs, the bias added to each agent's opinion will follow a normal distribution $\sim N(\mu, \sigma)$ with $\sigma = 0.03$ and $\mu = \pm 0.05$ to keep the opinion update in ranges $[-0.13, -0.06]$ and $[0.06, 0.13]$ and see how both affect opinion evolution.

Results

For both $\mu = -0.05$ and $\mu = +0.05$ we obtain:

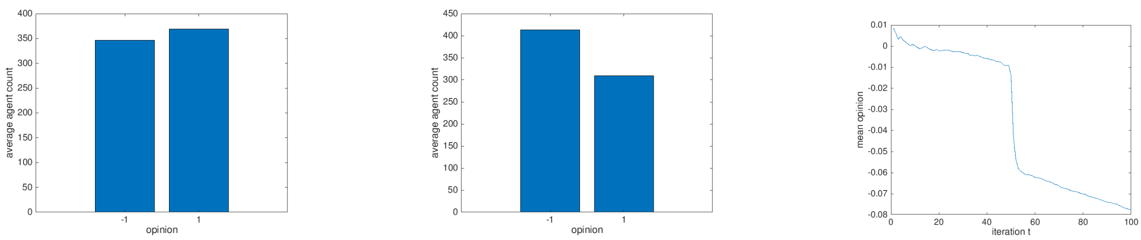


Figure 4.15: Referendum results at t_0 and t_{1000} and mean opinion evolution, $\mu = -0.05$

Discussion

We can now discuss the effect of the two type of events when opinion change is specific to each agent. Given that in both cases opinion was decreasing between t_0 and t_500 we can observe that while the first event event introducing bias of $\mu = -0.05$ seems to have precipitated opinion in

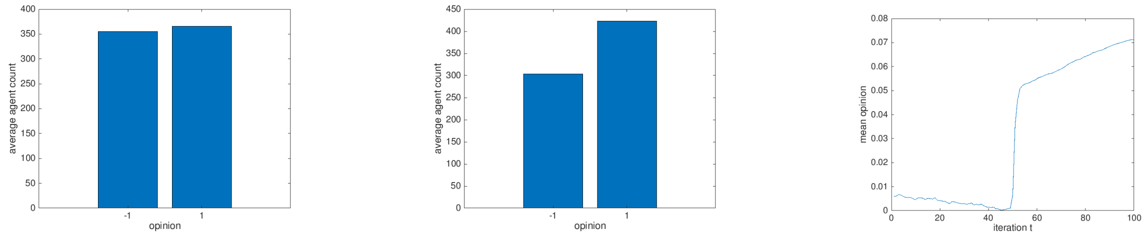


Figure 4.16: Referendum results at t_0 and t_{1000} and mean opinion evolution, $\mu = +0.05$

the direction of "remain", it appears that $\mu = +0.05$ has inverted this tendency as the opinion starting increasing after the opinion change. This is easily explainable given our model settings: shifting the opinion in a specific direction increases the number of stubborn agents in that direction while agents from the other side of the spectrum become more moderate and lose some of their stubbornness. Depending on which side the event pushes opinion therefore either precipitates opinion evolution or reverses its trend.

4.11.3 Simulation 3: multiple random events

We now wish to see how multiple events occurring during a simulation can influence opinion making of the population. Events will occur on a random basis with 1% chance per iterations so that our of 1000 iterations about 10 events will take place. Each event will bias opinion of agents on an individual basis following normal distribution $\sim N(\mu, \sigma)$ where $\sigma = 0.05$ and $\mu \sim U(-0.05, 0.05)$ decided each time an event occurs, so that opinion shift will occur around a shared mean for all agents, in a range of $[-0.2, 0.2]$ in 99.72% of cases. We run 100 simulations with a population of 10000 agents split in 500 groups of 20 agents and study the metric averaged on these 100 simulations.

Results

	t_0	t_{1000}
average sample mean	0.0053	0.0078
average sample standard deviation	0.4219	1.707
remain	49.22%	49.86%
leave	49.86%	50.14%

Table 4.40: Metrics over 100 simulations

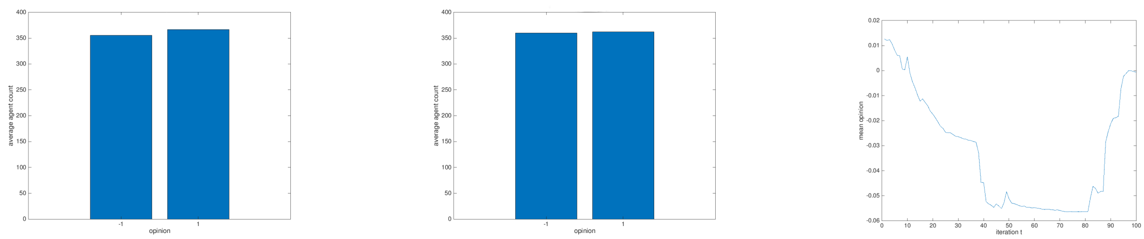


Figure 4.17: Referendum results at t_0 and t_{1000} and mean opinion evolution

Discussion

Comparing the results with running the model without events (table 4.37) it appears that introducing events does not induce any major changes in terms of opinion distribution or extremism. Referendums yield similar results as well. We can however highlight that final standard deviation is ~ 1.7 as we forgot to add opinion correction after events occurred in the model at the time of running the simulations (meaning that events could shift agents opinion beyond $[-1, 1]$ range) but was corrected before running final evaluations. Observing the mean opinion curve, we can see that while they do not appear to change opinion distribution, events introduce randomness in how evolution occurs.

Chapter 5

Testing

5.1 Evaluation

In order to make the model as realistic as possible, starting parameters were set to match the UK population at the time of the Brexit vote. The evaluation procedure will consist in evaluating how well it can reproduce the EU Referendum at t_0 before assessing how close the opinion evolution between t_0 and t_{1000} fits the polls conducted since June 23^{rd} 2016. The initial parameters will be set following three different distributions: the UK population (on which the model was build), the West Midlands population and the Scottish population. For all three evaluations, parameters will be set as follows:

parameter	value
population N	100,000
groups g	5,000
agents per group p	20
iterations t_{max}	1000

Table 5.1: Parameters used for the model

Initial distribution will be set to $\sim N(0, \frac{1}{3})$ with bias set according to the described distributions. Turnout is set to 72.2% and indecisiveness to 0.05. Events will occur with a 1% probability at each iteration, parametrised to shift opinion as described in Section 4.11. Finally, we will attempt to forecast how opinion has varied in Greater London since the 2016 EU Referendum and make predictions on which outcome a 2^{nd} referendum would have in London if it was ever to take place.

5.1.1 Modelling Great Britain

The first evaluation will consist in running the model implemented so far with initial opinion distribution following Section 4.9 description.

Results and discussion

	remain	leave
EU referendum results [2]	48.11%	51.89%
model results at t_0	49.16%	50.84%
model results at t_{max}	52.42%	47.58%

Table 5.2: Results: modelling Great Britain - 1000 iterations

We first observe that the initial distribution is accurate at $\pm 1\%$ when compared with EU Referendum results. While this is not an surprise per say as the model was designed so that its initial opinion distribution would fit accurately the referendum results, it is still worth noting that we have been able to obtain satisfying results.
We will now analyze how well we are able to model opinion evolution after the referendum.

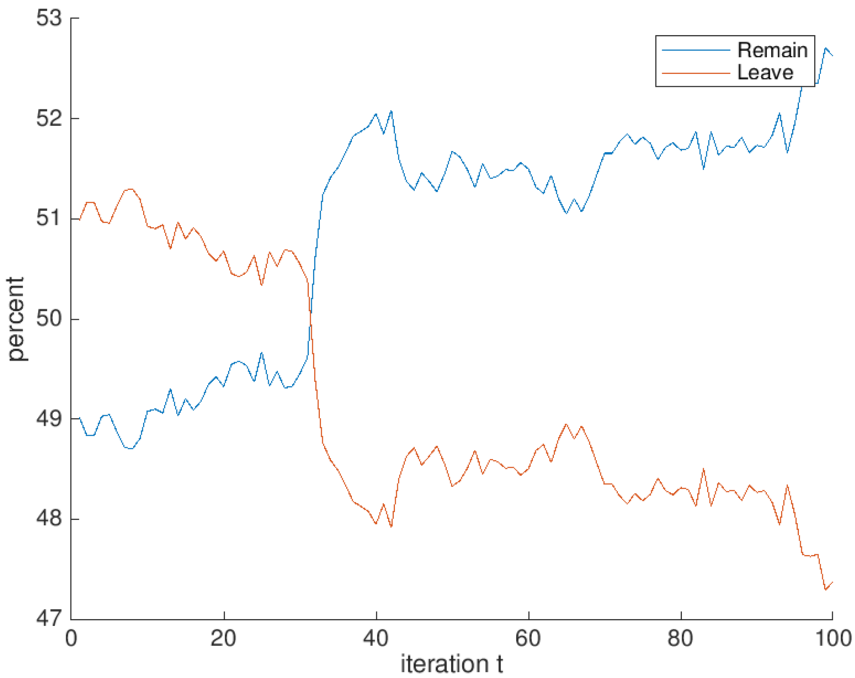


Figure 5.1: Model opinion evolution over 1000 iterations

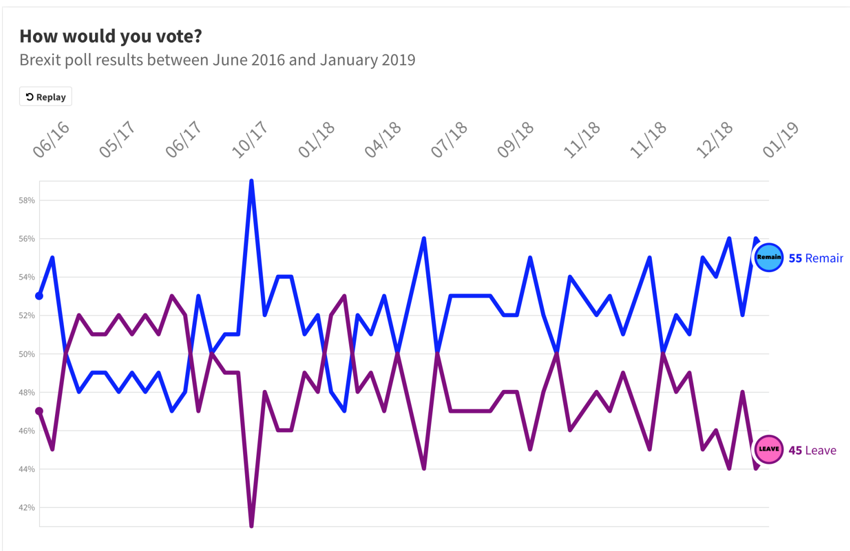


Figure 5.2: Poll conducted after the referendum [59]

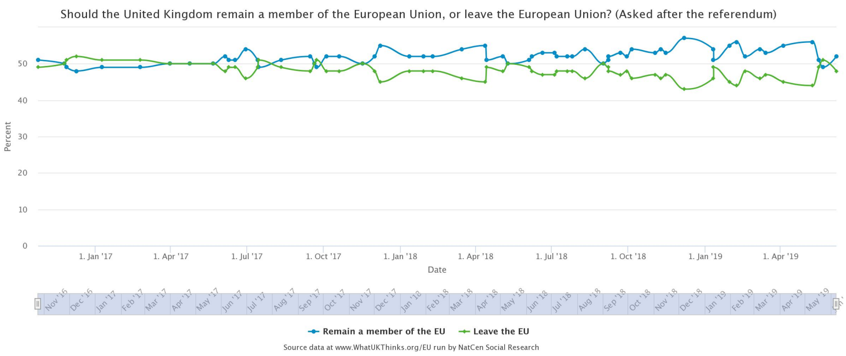


Figure 5.3: Poll conducted after the referendum [60]

We must first stress that our model ran on 1000 iterations, a metric that has not been associated with any real-life time measure. It is therefore pointless to claim that we have been able to model perfectly opinion evolution over a certain time-frame. We can however highlight that we have been able to model realistically opinion evolution, as our model predicts a polarization in population's opinion with the "remain" side taking over "leave", which is how opinion evolves in both poll collections.

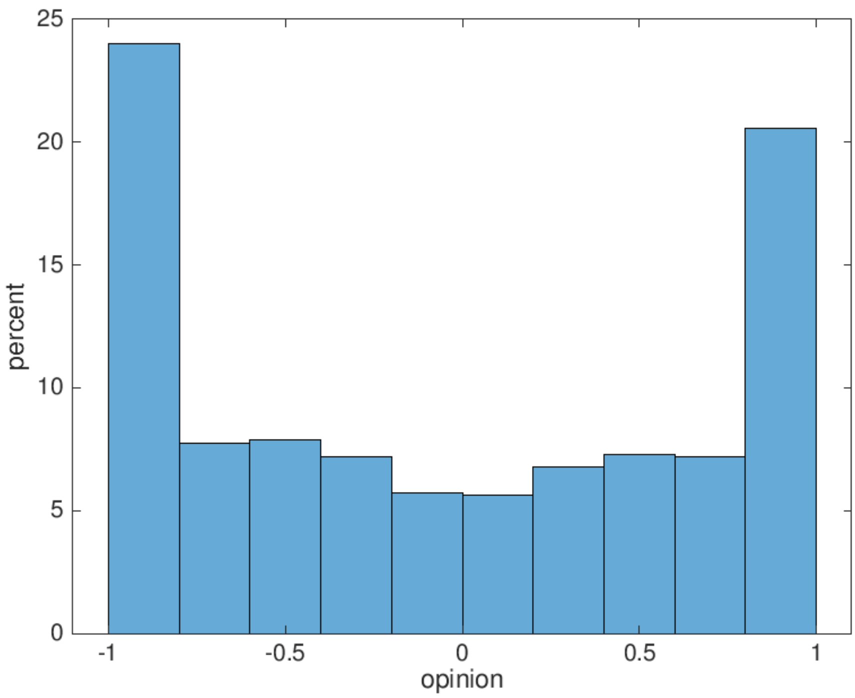


Figure 5.4: Model opinion distribution after 1000 iterations

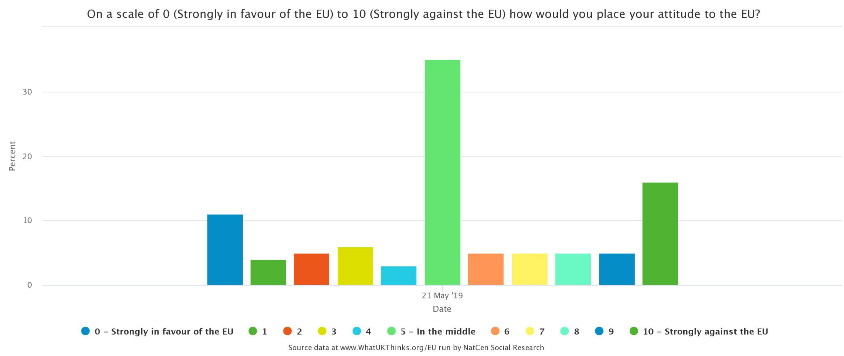


Figure 5.5: Poll conducted in May 2019: "On a scale from 1 to 10 how would you place your attitude to the EU?" [61]

In terms of opinion distribution, our model is been able to reproduce how opinion tends to cluster at the end of the distribution. We however see a dip in our model in the middle of the distribution, while in the poll the majority of agents are concentrated in this region of the spectrum. These two results can be traced back to our original assumption on opinion and stubbornness (see Section 4.3): agents with extreme opinion will be more stubborn while neutral agents will be more naive. While our assumption appears to be right regarding individuals with strong opinion, it is obviously false when it comes to agents in the middle of the spectrum. In our model, stubborn agents do not update their opinion with time and drag toward them neutral agents. In real-life, it appears that opinion clusters at the tail and center of the distribution. Had this poll been released earlier, we could have incorporated its results in our model.

5.1.2 Modelling West Midlands

age	population proportion	remain	leave	bias
18-24	11.5%	73%	27%	-0.23
25-34	15.21%	62%	38%	-0.12
35-44	18.44%	52%	48%	-0.02
45-54	16.45%	43%	57%	+0.07
55-64	15.89%	43%	57%	+0.07
65+	22.5%	40%	60%	+0.1

Table 5.3: West Midlands age distribution [39] and corresponding voting intention [42]

household income (£)	population proportion	remain	leave	bias
< 22000	50%	38%	62%	+0.12
> 22000	50%	65%	35%	0

Table 5.4: West Midlands household income distribution [39,62] and corresponding voting intention [42]

education level	population proportion	remain	leave	bias
higher education achieved	23%	68%	32%	-0.18
higher education not achieved	77%	41%	59%	+0.09

Table 5.5: West Midlands education level [63] and corresponding voting intention [42]

Results and discussion

	remain	leave
EU referendum results in West Midlands [64]	40.7%	59.3%
model results at t_0	35.09%	64.91%
model results at t_{max}	34.11%	65.89%

Table 5.6: Results: modelling West Midlands - 1000 iterations

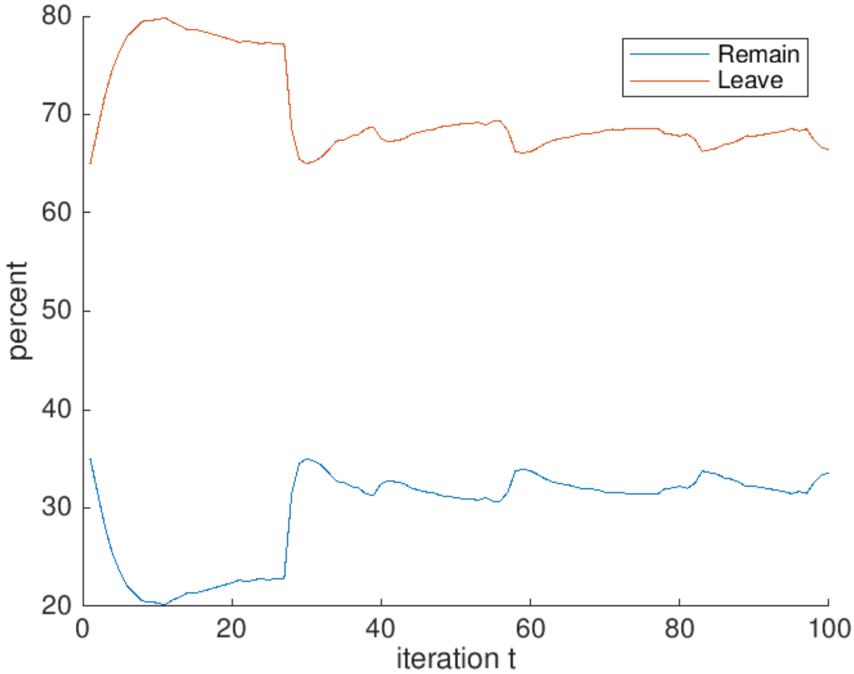


Figure 5.6: Model opinion evolution over 1000 iterations

At initialization, our model slightly overestimates the gap in opinion. The opinion evolution then appears to follow the trends of opinion observed recently in the region. We can refer to the recent European Election voting results which show a clear majority for Brexit in the West Midlands [65].

5.1.3 Modelling Scotland: model limits

We now evaluate how well we are able to model Scotland's opinion since the EU Referendum, using the following initial distribution parameters:

age	population proportion	remain	leave	bias
18-24	16.74%	73%	27%	-0.23
25-34	16.10%	62%	38%	-0.12
35-44	17.69%	52%	48%	-0.02
45-54	18.96%	43%	57%	+0.07
55-64	8.87%	43%	57%	+0.07
65+	21.64%	40%	60%	+0.1

Table 5.7: Scotland age distribution [39] and corresponding voting intention [42]

household income (£)	population proportion	remain	leave	bias
< 20000	41%	38%	62%	+0.12
$\in [20000, 40000[$	37%	47%	53%	+0.03
> 40000	22%	65%	35%	-0.15

Table 5.8: Scotland household income distribution [66, 67] and corresponding voting intention [42]

education level	population proportion	remain	leave	bias
higher education achieved	48%	68%	32%	-0.18
higher education not achieved	52%	41%	59%	+0.09

Table 5.9: Scottish education level [68] and corresponding voting intention [42]

	remain	leave
EU referendum results in Scotland [64]	62.0%	38.0%
model results at t_0	50.62%	49.38%
model results at t_{max}	51.04%	48.96%

Table 5.10: Results: modelling Scotland - 1000 iterations

Results and discussion

The model behaves poorly when attempting to model the initial opinion distribution based on Scotland population, as it yields a ratio well too balanced compared to the actual results for the referendum in Scotland. A reason for this is that our model relies too much on socio-demographic factors to distribute initial opinion, and does not account for political motivations of agents. That is, two regions can hold different beliefs regardless of their socio-demographics due to cultural beliefs and their respective history (which we did not even consider). Scotland expressed their will to stay in the United Kingdom back in 2014 with a very high ratio of 55 : 45 [69] which could have been accounted for in the model, assuming people's opinion in 2014 was also reflective of their thoughts about the European Union.
We now observe opinion evolution and compare it to polls conducted in Scotland:

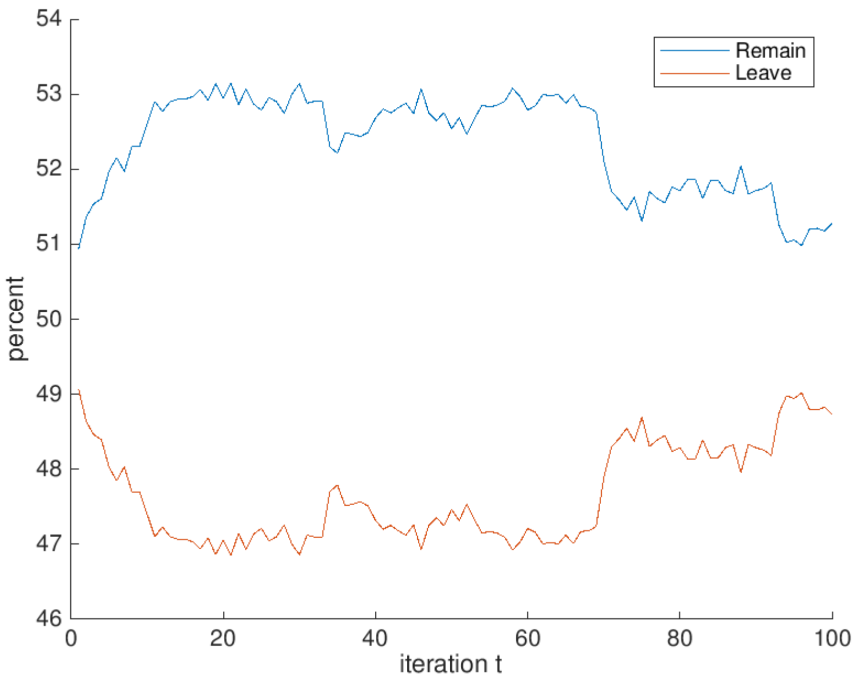


Figure 5.7: Model opinion evolution over 1000 iterations

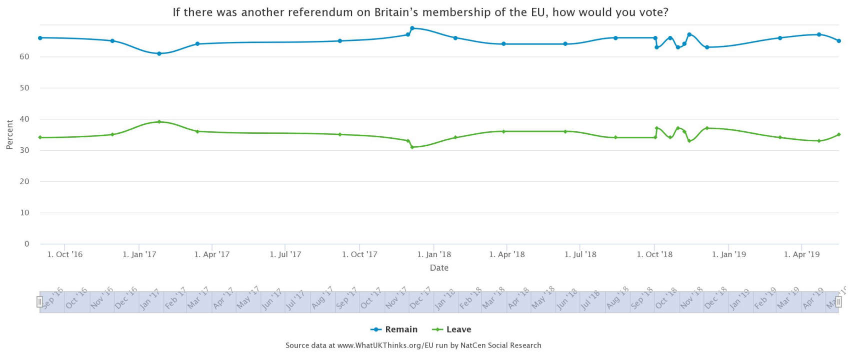


Figure 5.8: Polls conducted after the referendum [70]

Although opinion distribution at initialization is inexact, opinion evolution follows the trend observed in polls. A more exact initial distribution could have allowed for a more accurate opinion evolution modelling in this scenario.

5.1.4 Modelling Greater London

We end this project by modelling the opinion in London in an attempt to forecast how Londoners would vote in a second EU referendum.
We set the initial distribution with the following parameters:

age	population proportion	remain	leave	bias
18-24	10.86%	73%	27%	-0.23
25-34	24.74%	62%	38%	-0.12
35-44	19.94%	52%	48%	-0.02
45-54	17.15%	43%	57%	+0.07
55-64	12.46%	43%	57%	+0.07
65+	14.85%	40%	60%	+0.1

Table 5.11: Greater London age distribution [71] and corresponding voting intention [42]

household income (£)	population proportion	remain	leave	bias
< 20000	22.7%	38%	62%	+0.12
$\in [20000, 40000[$	38.48%	47%	53%	+0.03
$\in [40000, 65000[$	24.98%	58%	42%	-0.08
> 65000	13.84%	65%	35%	-0.15

Table 5.12: Greater London household income distribution [72] and corresponding voting intention [42]

education level	population proportion	remain	leave	bias
higher education achieved	70%	68%	32%	-0.18
higher education not achieved	30%	41%	59%	+0.09

Table 5.13: Greater London Education level [73, 74] and corresponding voting intention [42]

Results and discussion

The following results were obtained:

	remain	leave
EU referendum results [64]	59.9%	41.1%
model results at t_0	63.41%	46.59%
model results at t_{max}	62.54%	47.46%

Table 5.14: Results: modelling Great Britain - 1000 iterations

At initialization our model gets results in the range of the referendum results, which suggests that our model is at least applicable to model opinion distribution in London.
We now plot the predicted evolution of opinion:

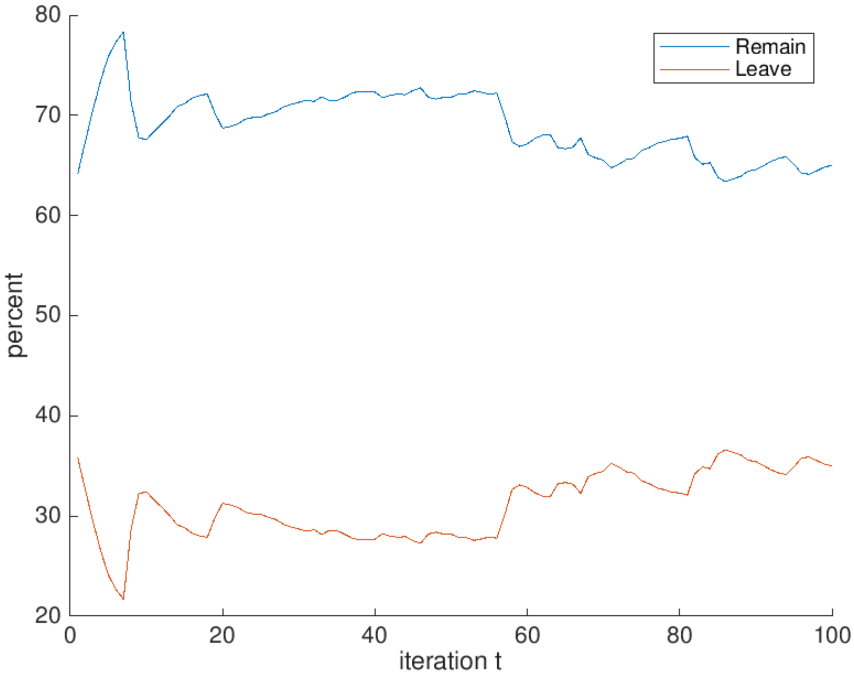


Figure 5.9: Forecasting opinion evolution in London

According to our model, if a second EU Referendum was to take place, the results in Greater London should be very similar to those of the first referendum.

Chapter 6

Conclusion and future work

6.1 Discussion and self-critique

6.1.1 Discussing the model

From the results obtained during the evaluation procedure, we can first conclude that the initial opinion distribution relies too much on socio-demographic factors. While this provided accurate results in the overall UK population West Midlands and London, the results did not match the Scottish referendum outcome.
Regarding opinion evolution, in all three cases we evaluated, the model yielded results which followed the trends observed in real-life scenario. These results were however very dependent upon initial distribution, as observed for the Scottish population modelling. In conclusion, the model is able to get satisfying results

6.1.2 Methodology

The employed methodology was successful in the sense that adopting an iterative and incremental approach allowed for gradual improvement of the project and constant evaluation of the model. This allowed for a smooth development process which was quite motivating as each new increment yielded better results. It also allowed to follow and document closely the development process. Alternating between initial opinion distribution and opinion evolution improvements was maybe not ideal, and a better approach could be to focus first on opinion distribution (Sections 4.4, 4.5, 4.37) before moving on to opinion dynamics (Sections 5.1:3, 5.6:8 and 5.10:11). Evaluation metrics should also have been defined from the very beginning rather than considered at each new improvements, which would have allowed to avoid several mistakes - see Section 4.11.1 where it appeared that convergence rate was erroneous when introducing a new way of measuring opinion evolution - and make a more consistent evaluation of the project's development.

6.1.3 Assumptions and improvements

The model was developed following several assumptions and using arbitrary parameters which could have been explored and studied with more depth.
To start with, improvements were added and tested on a low population of 1000 agents, a small sample which should have been increased to at least 10,000 earlier (it was only done when introducing events, see Section 4.11).
Iterations were set to 1000 arbitrarily, as it intuitively would allow agents to interact a finite num-

ber of time and have on average $\frac{1}{10} \times 1000 = 100$ interactions (2 random agents interact per group of 20 at each iteration). No attempt to link iterations with real-life time-frame was made, but it could have provided a clearer insight of the model performance when comparing it with polls during evaluation. Future models should investigate this question as it would allow for better interpretations and more accurate improvements.

6.2 Future work: other interesting ideas

In addition to the improvements just discussed, further enhancements are described in this section.

6.2.1 Socio-demographics

Several socio-cultural factors, if taken into account, can provide a more comprehensive view of the problem of opinion dynamics and enhance the model's complexity and realism in the context of the EU Referendum.

Ethnic origins

Many arguments of the Brexit campaign involved European and international immigration, and ethnics have been shown to influence individual's vote outcome [40,75]. Combining research insights with the 2011 census statistics [39] which include ethnics data, it could be possible to include this feature into a model.

Group formation

Our model formed groups randomly at initialisation. As mentioned in the discussion of Section 4.4, a more realistic model should account for the fact that social groups usually display some kind of similitude between individuals, whether it is ethnicity, age or social-class.Further improvements can therefore focus on rectifying group formation by taking into account opinion of agents.

Urban and rural area

Urban and rural areas usually vary substantially in terms of population density and interactions between people. In a more dense population, we would expect people to interact more frequently than in a sparsely populated area. Furthermore, we can intuitively expect citizens to have more "superficial" interactions in social situations while individuals in rural areas should tend to know better those they interact with (these assumptions would of course need to be verified). Differentiating rural and urban areas could therefore add realism to the model.

6.2.2 Media

Both traditional and digital media are important sources of information for individuals in our society [76,77], with substantial influence on decision formation in the case of the Brexit referendum [40], and their role should be taken into account if one wishes to achieve a comprehensive model of opinion dynamics in the light of a political decision.

Possible extension projects

Traditional media's impact (television, radio, press) in today's society is still a subject of debate among researchers, others supporting it is only a minor source of influence [78]. Regardless of the

final conclusion, we propose the following method to incorporate press publications in an opinion dynamics model before a referendum, applied to the case-study of Brexit.
A possible starting point could be to reproduce Moore et al. [79]'s work on media influence in networks under Deffuant interactions and make it specific to the Brexit referendum scenario studied in this project.
Going further and adopting a different approach, with the help of a web-scrapping script, one could collect all press-articles published online prior to the referendum. Using machine learning techniques, it is then possible to determine with high precision whether publications are pro or con Brexit. The final step is to correlate the articles' alignment with polled opinion and conclude how strong the relationship is, which will then allow for implementation in the model.

Social networks

Several studies support the fact that social media plays an increasing role on public information and opinion formation in the case of the Brexit referendum [80, 81], shaping public opinion and leading to formation of echo chambers reinforcing opinions [81]. Similarly to traditional media, sentiment analysis techniques can allow for correlating online opinion to real-life polling and ultimately include social media interactions into a comprehensive model of opinion dynamics.

6.2.3 Employed agents

The impact of employed agents [36], described in section 2.5.1, could also be added to the model. Employed agents are individuals employed by a political party to influence the opinion of public individuals. They can be easily modelled as agents of maximum stubbornness (who will never update their own opinion) but will still interact with other non-stubborn agents who will update their opinion accordingly, leading to a unilateral shift in opinion.

6.2.4 Discrete opinion model

Our model was representing individuals' opinions as a continuous variable ranging from -1 to +1. A variation could be to represent opinion as a discrete variable grouping opinion into "blocks". That is, extreme agents could have opinion valued as ± 1 while indecisive or moderate individuals would have an opinion of 0. Extension could be achieved by adding on variable range: $[-3, -2, -1, 0, 1, 2, 3]$ each discrete variable accounting for a degree of strength in the agent's opinion, allowing to group individuals according to their position more easily.

List of Figures

2.1 Three round voting hierarchy with group size $r = 3$, from *Sociophysics: a review of Galam models*" [15] . . . 12
2.2 Example of a weighted network of individuals from *Introduction to Abstract Algebra with Applications to Social Systems* lecture by Dr. Arceneaux at Princeton . . . 15
2.3 Pitchfork bifurcation of opinion, consensus reaching depending on stop-signal rate, from Naomi Leonard's lecture *Learning from Nature: When Models Meet Multitude* [33] . . . 17

3.1 High level view - two sub-groups each containing five individuals . . . 19
3.2 High level view - agents interacting inside sub-groups . . . 20
3.3 High level view - social mobility . . . 20
3.4 ULM class diagram of Individual (made with draw.io) . . . 22

4.1 Low bound $d = 0.25$ . . . 27
4.2 Low bound $d = 0.25$ inner-group opinion distribution after 1000 iterations . . . 27
4.3 Low bound $d = 0.25$ . . . 28
4.4 High bound $d = 1.0$ inner-group opinion distribution after 1000 iterations . . . 29
4.5 Initial stubbornness function . . . 30
4.6 Agents' opinions within two groups after 1000 iterations . . . 34
4.7 $s_1 = 2 \times \mid opinion \mid$ . . . 38
4.8 $s_2 = 2 \times \mid \sqrt{opinion} \mid$ . . . 38
4.9 $s_3 = 2.63 \times \mid tanh(opinion) \mid$ . . . 38
4.10 $s_4 = 2.57 \times \mid arctan(opinion) \mid$ . . . 38
4.11 Comparing referendum results averaged over 100 simulations for the three models at t_{1000}: indecisiveness, turnout and combined (turnout changes the scale as 27.8% of agents were removed) . . . 42
4.12 Francois Hollande's public approval [58] . . . 43
4.13 Mean opinion evolution when incorporating event with gravity 0.05 . . . 44
4.14 Mean opinion evolution when incorporating event with gravity 0.05 and new convergence rate rule . . . 45
4.15 Referendum results at t_0 and t_{1000} and mean opinion evolution, $\mu = -0.05$ . . . 45
4.16 Referendum results at t_0 and t_{1000} and mean opinion evolution, $\mu = +0.05$ . . . 46
4.17 Referendum results at t_0 and t_{1000} and mean opinion evolution . . . 46

5.1 Model opinion evolution over 1000 iterations . . . 49
5.2 Poll conducted after the referendum [59] . . . 49
5.3 Poll conducted after the referendum [60] . . . 50
5.4 Model opinion distribution after 1000 iterations . . . 50

5.5 Poll conducted in May 2019: "On a scale from 1 to 10 how would you place your attitude to the EU?" [61] 51
5.6 Model opinion evolution over 1000 iterations 52
5.7 Model opinion evolution over 1000 iterations 53
5.8 Polls conducted after the referendum [70] 54
5.9 Forecasting opinion evolution in London 55

A.1 Reference bound $d = 0.5$ 68
A.2 High bound $d = 0.75$ 68
A.3 Low confidence bound $d = 0.25$ - wide opinion spread 68
A.4 Mesh 68
A.5 Histogram after 800 iterations 68
A.6 Histogram after 1600 iterations 68
A.7 Low confidence bound $d = 0.25$ - low opinion spread 69
A.8 Mesh 69
A.9 Histogram after 800 iterations 69
A.10 Histogram after 1600 iterations 69
A.11 Low convergence rate $\mu = 0.1$ 69
A.12 Mesh 69
A.13 Histogram after 800 iterations 69
A.14 Histogram after 1600 iterations 69
A.15 High convergence rate $\mu = 0.75$ 70
A.16 Mesh 70
A.17 Histogram after 800 iterations 70
A.18 Histogram after 1600 iterations 70

Bibliography

[1] Eu referendum timeline: Countdown to the vote. https://www.bbc.co.uk/news/uk-politics-33141819. Accessed: 2019-04-06.

[2] Eu referendum results. https://www.electoralcommission.org.uk/find-information-by-subject/elections-and-referendums/past-elections-and-referendums/eu-referendum/electorate-and-count-information. Accessed: 2019-03-06.

[3] R. S. Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2):175–220, 1998.

[4] A. Koriat, S. Lichtenstein, and B. Fischhoff. Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2):107–119, 1980.

[5] A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131, 1974.

[6] A. Furnhama and H. C. Boo. A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40(1):35–42, 2011.

[7] H. M. Johnson and C. M. Seifert. Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology Learning Memory and Cognition*, 20(6):1420–1436, 1994.

[8] I. L. Janis. *Victims of Groupthink: A psychological study of foreign-policy decisions and fiascoesk*. Houghton, Mifflin, 1 edition, 1972.

[9] S. E. Asch. Effects of group pressure upon the modification and distortion of judgments. *Groups, leadership and men; research in human relations*, page 177–190, 1951.

[10] E. Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, 1925.

[11] S. Galam, Y. Gefen, and Y. Shapir. Sociophysics: A new approach of sociological collective behaviour. i. mean behaviour description of a strike. *The Journal of Mathematical Sociology*, 9(1):1–13, 1982.

[12] P. Clifford and A. Sudbury. A model for spatial conflict. *Biometrika*, 60(3):581–588, 1973.

[13] S. Galam. Minority opinion spreading in random geometry. *The European Physical Journal B*, 25:403–406, 2002.

[14] S. Galam. Social paradoxes of majority rule voting and renormalization group. *Journal of Statistical Physics*, 61(3–4):943–951, 1990.

[15] S. Galam. Sociophysics: A review of galam models. *International Journal of Modern Physics C*, 19(3):409–440, 2008.

[16] S. Galame. Le dangereux seuil critique du fn. Le Monde, 5 1997. French newspaper.

[17] S. Galame. Crier, mais pourquoi. Libération, 4 1997. French newspaper.

[18] French Government. Résultats du 1er tour. French Gov. Website, 4 2002. https://www.conseil-constitutionnel.fr/election-presidentielle-2002/resultats-du-1er-tour.

[19] French Government. Résultats définitifs du second tour - 5 mai 2002. French Gov. Website, 5 2002. https://www.conseil-constitutionnel.fr/election-presidentielle-2002/resultats-definitifs-du-second-tour-5-mai-2002.

[20] K. Sznajd-Weron and J. Sznajd. Opinion evolution in closed community. *International Journal of Modern Physics C*, 11(6):1157–1165, 2001.

[21] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(4):87–98, 2000.

[22] M. H. DeGroot. Reaching a consensus. *Journal of American Statistical Association*, 69(345):118–121, 1974.

[23] T. A. Arceneaux. Introduction to abstract algebra with applications to social systems. lecture notes, 2010. Princeton University.

[24] N. E. Friedkin and E. C. Johnsen. Social influence and opinions. *Journal of Mathematical Sociology*, 15(3-4):193–205, 1990.

[25] N. E. Friedkin and E. C. Johnsen. Social influence networks and opinion change. *Advances in Group Processes*, 16:1–29, 1999.

[26] D. Grünbaum. Schooling as a strategy for taxis in a noisy environment. *Evolutionary Ecology*, 12(5):503–522, 1998.

[27] I. D. Couzinnw, J. Krausew, R. Jamesz, G. D. Ruxtony, and Nigel R. Franksz. Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, 218:1–11, 2002.

[28] J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. J. Hale, E. Despland, E. R. Miller, and S. J. Simpson. From disorder to order in marching locusts. *Science*, 302(5778):1402–1406, 2006.

[29] K. Von Frisch. *Bees.* Cornell University Press, 1950.

[30] T. D. Seeley, P. K. Visscher, T. Schlegel, P. M. Hogan, N. R. Franks, and J. A. R. Marshall. Stop signals provide cross inhibition in collective decision-making by honeybee swarms. *Science*, 335(6065):108–111, 2012.

[31] D. Pais, P. M. Hogan, T. Schlegel, N. R. Franks, and N. E. Leonard. A mechanism for value-sensitive decision making. *PLOS One*, 8(9), 2013.

[32] N. E. Leonard. Learning from nature: When models meet multitude. IEEE Conference on Decision and Control, Melbourne, Australia, 12 2017.

[33] N. E. Leonard. Learning from nature: When models meet multitude. lecture notes, 12 2017. Princeton University.

[34] C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *International Journal of Modern Physics C*, 81(2):591–646, 2009.

[35] A. V. Proskurnikova and R. Tempo. A tutorial on modeling and analysis of dynamic social networks. part 1. *Annual Reviews in Control*, 43:65–79, 2017.

[36] S. Brugger and C. Schwirzer. Opinion formation by "employed agents" in social networks. How it was published, 5 2011. Lecture with Computer Exercises: Modelling and Simulating Social Systems with MATLAB - Student Project Report - Swiss Federal Institute of Technology Zurich.

[37] C. Letta and V. Mirabet. Modelling the dynamics of animal groups in motion. *South African Journal of Science*, 104:192–198, 2008.

[38] M. Moussaïd, J. E. Kämmer, P. P. Analytis, and H. Neth. Social influence and the collective dynamics of opinion formation. *PLoS One*, 8(11), 2013.

[39] Population of england and wales. `https://www.ons.gov.uk/census/2011census/2011censusdata/2011censusdatacatalogue/populationstatisticsfortheuk`. Accessed: 2019-03-06.

[40] S. Becker, T. Fetzer, and D. Novy. Who voted for brexit? a comprehensive district-level analysis. *Economic Policy*, 32(92):601–650, 2017.

[41] A. Zhang. New findings on key factors influencing the uk's referendum on leaving the eu. *World Development*, 102(2018):304–314, 2017.

[42] Yougov brexit survey 23-24 june 2016. `https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/oxmidrr5wh/EUFinalCall_Reweighted.pdf`. Accessed: 2019-13-06.

[43] Ipsos: How britain voted in the 2016 eu referendum. `https://www.ipsos.com/ipsos-mori/en-uk/how-britain-voted-2016-eu-referendum`. Accessed: 2019-13-06.

[44] M. Goodwin and O. Heath. *Brexit vote explained: poverty, low skills, and lack of opportunity.* Joseph Rowntree Foundation, 2016.

[45] Most people in the uk do not go to university – and maybe never will. `https://www.theguardian.com/higher-education-network/blog/2013/jun/04/higher-education-participation-data-analysis`. Accessed: 2019-13-06.

[46] The effects of taxes and benefits on household income, disposable income estimate: 2016 (ons). `https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/datasets/householddisposableincomeandinequality`. Accessed: 2019-13-06.

[47] Politico: Hypothetical 2nd brexit referendum. `https://www.politico.eu/2019-european-elections/united-kingdom/`. Accessed: 2019-14-06.

[48] Yougov / times survey results 12-13/06/16. `https://d25d2506sfb94s.cloudfront.net/cumulus_uploads/document/0q7lmn19of/TimesResults_160613_EUReferendum_W_Headline.pdf`. Accessed: 2019-14-06.

[49] File:uk eu referendum polling.svg. `https://commons.wikimedia.org/wiki/File:UK_EU_referendum_polling.svg`. Accessed: 2019-14-06.

[50] M.J. Goodwin and O.Heath. The 2016 referendum, brexit and the left behind: An aggregate level analysis of the result. *The Political Quarterly*, 87(3):323–332, 2016.

[51] D. Manley, K. Jones, and R. Johnston. The geography of brexit – what geography? modelling and predicting the outcome across 380 local authorities. *Local Economy: The Journal of the Local Economy Policy Unit*, 32(3):183–203, 2017.

[52] Labour mp jo cox dies after being shot and stabbed. BBC, June 2016. https://www.bbc.co.uk/news/uk-england-36550304, Accessed: 2019-15-06.

[53] Labour mp jo cox dies after being shot and stabbed. The Guardian, June 2016. https://www.theguardian.com/uk-news/2016/jun/16/labour-mp-jo-cox-shot-in-west-yorkshire, Accessed: 2019-15-06.

[54] D. Dobreva, D. Grinnell, and M. Innes. Prophets and loss: How "soft facts" on social media influenced the brexit campaign and social reactions to the murder of jo cox mp. *Policy and Internet*, 2019.

[55] S. Gonzalez-Bailon, R. E. Banchs, and A. Kaltenbrunner. Emotional reactions and the pulse of public opinion: Measuring the impact of political events on the sentiment of online discussions. *CoRR*, abs/1009.4019, 2010.

[56] Explosion de la popularité de françois hollande dans les sondages. Le Monde, December 2015. https://www.lemonde.fr/politique/article/2015/12/01/hollande-conquiert-desormais-la-moitie-des-francais_4821824_823448.html, Accessed: 2019-15-06.

[57] La popularité de hollande en hausse après les attentats. Le Parisien, November 2015. http://www.leparisien.fr/politique/la-popularite-de-hollande-en-hausse-apres-les-attentats-24-11-2015-5307121.php, Accessed: 2019-15-06.

[58] Rick Noack. The reason french president hollande won't seek re-election, in one chart. The Washington Post, December 2016. https://www.washingtonpost.com/news/worldviews/wp/2016/12/01/the-reason-french-president-hollande-wont-seek-reelection-in-one-chart/?utm_term=.7444f8ecb278, Accessed: 2019-15-06.

[59] What uk thinkgs: Should the united kingdom remain a member of the european union, or leave the european union? (asked after the referendum). https://whatukthinks.org/eu/questions/should-the-united-kingdom-remain-a-member-of-the-european-union-or-leave-the-european-union-a?removed#. Accessed: 2019-17-06.

[60] The ferret: How would you vote? brexit poll results between june 2016 and january 2019. https://public.flourish.studio/visualisation/224833/?utm_source=showcase&utm_campaign=visualisation/224833. Accessed: 2019-17-06.

[61] What uk thinkgs: On a scale of 0 (strongly in favour of the eu) to 10 (strongly against the eu) how would you place your attitude to the eu? https://whatukthinks.org/eu/questions/on-a-scale-of-0-strongly-in-favour-of-the-eu-to-10-strongly-against-the-eu-how-would-you-pla#. Accessed: 2019-17-06.

[62] Ons: West midlands household distribution. https://www.ons.gov.uk/economy/regionalaccounts/grossdisposablehouseholdincome/bulletins/regionalgrossdisposablehouseholdincomegdhi/1997to2016. Accessed: 2019-17-06.

[63] West midlands european regional development fund operational programme (page 302). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/315877/WM_OP_FINAL_v3.1.pdf. Accessed: 2019-17-06.

[64] Bloomberg:eu referendum: Final results. https://www.bloomberg.com/graphics/2016-brexit-referendum/. Accessed: 2019-17-06.

[65] Bbc:european election 2019: Brexit party tops west midlands polls. https://www.bbc.co.uk/news/uk-england-48417494. Accessed: 2019-17-06.

[66] Scottish household survey 2017: annual report. https://www.gov.scot/publications/scotlands-people-annual-report-results-2017-scottish-household-survey/pages/3/. Accessed: 2019-17-06.

[67] Scottpho:income and employment: household income. https://www.scotpho.org.uk/life-circumstances/income-and-employment/data/household-income/. Accessed: 2019-17-06.

[68] Regional employment patterns in scotland: statistics from the annual population survey 2017. https://www.gov.scot/publications/regional-employment-patterns-scotland-statistics-annual-population-survey-2017/pages/6/. Accessed: 2019-17-06.

[69] Bbc:scottish referendum: Scotland votes 'no' to independence. https://www.bbc.co.uk/news/uk-scotland-29270441. Accessed: 2019-17-06.

[70] What uk thinks: If there was another referendum on britain's membership of the eu, how would you vote? (scotland). https://whatukthinks.org/eu/questions/if-there-was-another-referendum-on-britains-membership-of-the-eu-how-would-you-vote/?removed. Accessed: 2019-17-06.

[71] Trust for london: London's population by age. https://www.trustforlondon.org.uk/data/londons-population-age/. Accessed: 2019-14-06.

[72] Data gov: Focus on london - income and spending. https://data.gov.uk/dataset/f0e9be9a-2b59-4b8f-9009-91ce69a9c397/focus-on-london-income-and-spending. Accessed: 2019-14-06.

[73] Ons: Qualifications of working age population (nvq). https://data.london.gov.uk/dataset/qualifications-working-age-population-nvq-borough. Accessed: 2019-14-06.

[74] Eurostat: Tertiary educational attainment, age group 25-64 by sex and nuts 2 regions. https://ec.europa.eu/eurostat/tgm/graph.do?tab=graph&plugin=1&language=en&pcode=tgs00109&toolbox=type. Accessed: 2019-14-06.

[75] E. Alabrese, S.O. Beckerb, T. Fetzerc, and D. Novy. Who voted for brexit? individual and regional data combined. *European Journal of Political Economy*, 56:132–150, 2019.

[76] S. L. Althaus and D. Tewksbury. Patterns of internet and traditional news media use in a networked community. *Political Communication*, 17(1):21–45, 2000.

[77] M. J. Dutta-Bergman. Complementarity in consumption of news types across traditional and new media. *Journal of Broadcasting and Electronic Media*, 48(1):41–60, 2004.

[78] K. Newton. May the weak force be with you: The power of the mass media in modern politics. *European Journal of Policital Research*, 45(2):209–234, 2006.

[79] T. Moore, P. Finley, N. Brodsky, T. Brown, B. Apelberg, B. Ambrose, and R. Glass. Modelling education and advertising with opinion dynamics. *Journal of Artificial Societies and Social Simulation*, 18(2), 2015.

[80] M. Hänska and S. Bauchowitz. Tweeting for brexit: how social media influenced the referendum. *Brexit, Trump and the Media*, pages 31–35, 2017.

[81] M. Del Vicario, F. Zollo, G. Caldarelli, A. Scala, and W. Quattrociocchi. Mapping social dynamics on facebook: The brexit debate. *Social Networks*, 50:6–16, 2017.

[82] European union (withdrawal) act main motion (prime minister). `https://commonsvotes.digiminster.com/Divisions/Details/562`. Accessed: 2019-03-06.

Appendix A

Preliminary experiment - Deffuant model

A.1 Purpose

The aim of this first section was to run a simulation of the Deffuant model, varying parameters to get a comprehensive hands-on understanding of how to model a population's opinion dynamics. First experiments were ran with same parameters as described in section 2.3.1 to allow for comparisons with established models and results.

A.2 Running documented experiments

Running model following Deffuant's experiment

We now simulate the traditional Deffuant model as described in section 2.3.1. We will use a population of $N = 100$ agents which we store in a 1×100 vector. We set the parameters $d = 0.5$, $\mu = 0.5$. At each iteration t, two random actors meet and update their own opinion according to the Deffuant rule. At then build a $t \times 100$ matrix which will allow us to study the evolution of opinion iteration after iteration. We set initial opinions following a continuous distribution such that $X \sim U(0, 1)$.

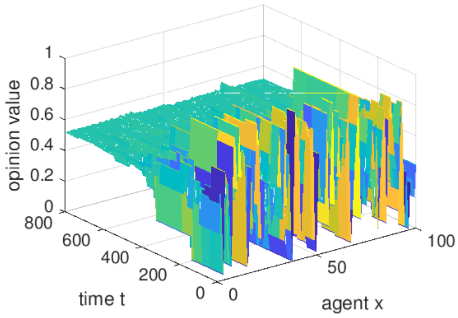
A.2.1 Experimenting

Varying d

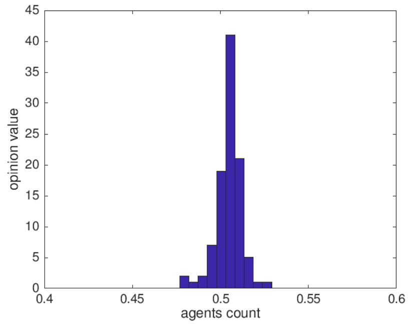
Setting $d = 0.5$, we observe that the population's opinion converges towards 0.5 at midpoint between the values 0 and 1 it could take. Our population reaches close-consensus after about $t = 600$ iterations as can be seen in figure A.1a. Final distribution can be seen in figure A.1b. Increasing d to $d = 0.75$, we observe a similar result of convergence towards consensus, but with a wider final distribution of opinion as shown in figure A.2b.

We then lower d to $d = 0.25$. The most noticeable effect is that agents do not reach consensus anymore, but rather form two distinct clusters of opinion around 25 and 75. The ratio between the two usually varies between $\sim 25 : 75$ (in either way) to $\sim 50 : 50$ (thus why we show both results in figures A.6 and A.10). Clustering takes more iterations to be reached and duration has been extended from $t_{max} = 800$ to $t_{max} = 1600$. As opinion is no longer uniform but spreading

Figure A.1: Reference bound $d = 0.5$

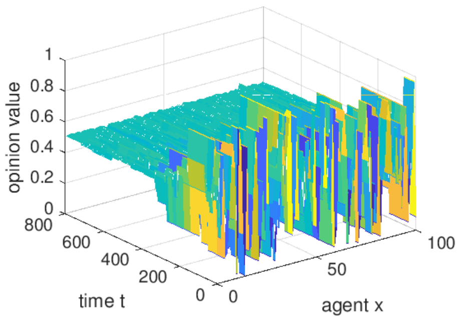


(a) Mesh

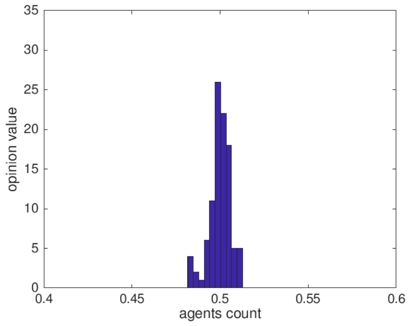


(b) Histogram after 800 iterations

Figure A.2: High bound $d = 0.75$



(a) Mesh



(b) Histogram after 800 iterations

between two options, 3D plots are harder to read and reader should refer mostly to histograms. Theses results are in agreement with those obtained by Deffuant (see section 2.3.1).

The Deffuant model with low confidence bound appears to be an interesting model for modelling population opinion evolution in a referendum situation as the end result shows a split in opinion between two options. The final distribution is furthermore close to what we usually observe in real scenarios, with very close to 50 : 50 result as in Brexit referendum to 30 : 70 as in the Parliament result.

Figure A.3: Low confidence bound $d = 0.25$ - wide opinion spread

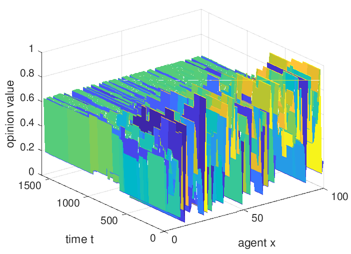


Figure A.4: Mesh

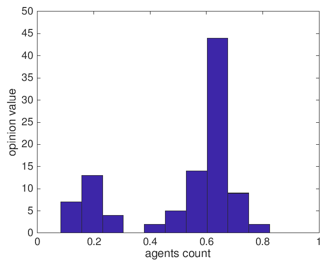


Figure A.5: Histogram after 800 iterations

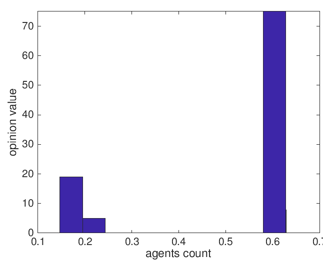


Figure A.6: Histogram after 1600 iterations

Figure A.7: Low confidence bound $d = 0.25$ - low opinion spread

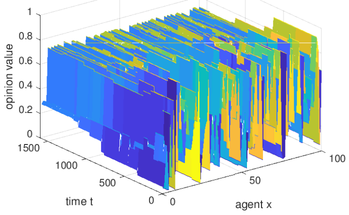


Figure A.8: Mesh

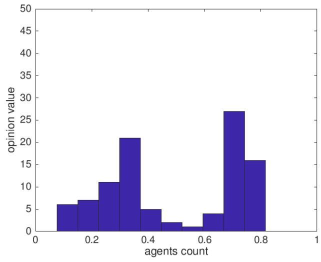


Figure A.9: Histogram after 800 iterations

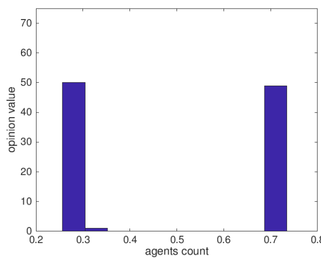


Figure A.10: Histogram after 1600 iterations

Varying μ

We now vary the convergence parameter μ to study how it affects the rate of opinion convergence. To obtain comparable results, we set $d = 0.5$ and run 1600 iterations observing the result at 800 and 1600. We will compare the results by studying how fast opinion approaches 0.5.

For $\mu = 0.1$, decreasing the convergence parameter we observe as expected that the model, while converging toward 0.5, does so more slowly as on both iteration 800 (see figures A.12, A.13) and 1600 (see figure A.14) we can observe that the opinion is still more widely spread than when $\mu = 0.5$ with the same confidence bound $d = 0.5$.

Figure A.11: Low convergence rate $\mu = 0.1$

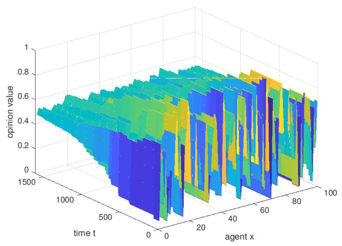


Figure A.12: Mesh

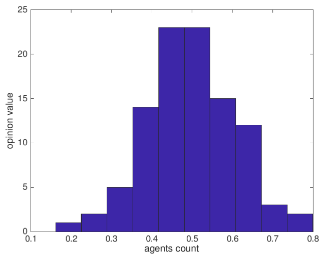


Figure A.13: Histogram after 800 iterations

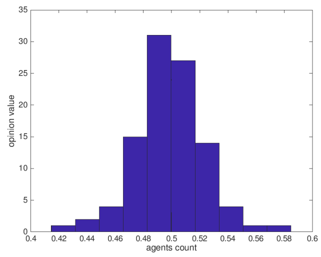


Figure A.14: Histogram after 1600 iterations

For $\mu = 0.75$, one could expect the model to converge faster than when $\mu = 0.5$ since the convergence parameter increases. However, we notice on iteration 800 that while the model for $\mu = 0.5$ had converged, the model with $\mu = 0.75$ still has more widely spread opinion (see figures A.16 and A.17) until finally reaching consensus at about 1000 iterations (see figure A.18).

We also notice at iteration 800 in both the histogram and 3D plot that we have an outlier agent (fig. A.16 and A.17) which disappears when consensus is reached at iteration 1000 (see fig. **??** and A.18) illustrating clearly how consensus is reached under large confidence bound $d = 0.5$ as outliers ultimately meet agents from the majority group and have their opinion converge towards theirs.

First conclusions

Regarding these first experiments, we obtain results which confirm all described simulations of the Deffuant model. We can conclusively say that our modelling works as expected and move on to the next stage. Furthermore, parallels can be drawn from our simulations and real scenarios.

Figure A.15: High convergence rate $\mu = 0.75$

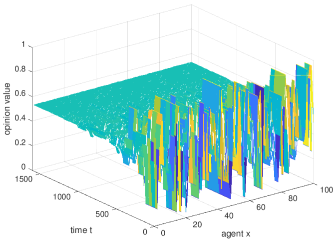


Figure A.16: Mesh

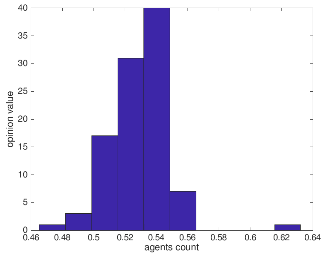


Figure A.17: Histogram after 800 iterations

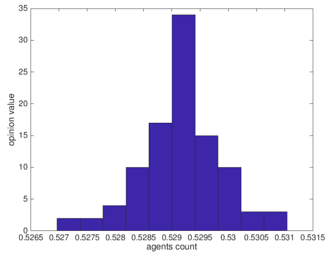


Figure A.18: Histogram after 1600 iterations

Assuming our population has initial opinion continuously distributed at t_0, that is when referendum is announced, we usually observe two situations: either a final result with a close to 50 : 50 ratio as in the EU referendum on June 23^{rd}, 2016 (which was 51.9 : 48.1 in favour of leave) [2] or a situation with wider gap in opinion with a 70 : 30 ratio as in the Parliament votes for Brexit deal on January 15^{th}, 2019 (which was 32 : 68 in favour of noes) [82]. If using a confidence bound when building our model, we should aim at tuning it according to the population modelled, which seems to be different whether they are a population (low confidence bound) or politics (high confidence bound).